

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

Applicant: Motoki KATO

International Application No.: PCT/JP2004/005782

International Filing Date: April 22, 2004

For: INFORMATION PROCESSING DEVICE,  
INFORMATION PROCESSING METHOD, PROGRAM,  
AND RECORDING MEDIUM

745 Fifth Avenue  
New York, NY 10151

**EXPRESS MAIL**

Mailing Label Number: EV206809644US

Date of Deposit: January 7, 2005

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" Service under 37 CFR 1.10 on the date indicated above and is addressed to Mail Stop PCT, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

Adam Ahmed  
(Typed or printed name of person mailing paper or fee)

[Signature]  
(Signature of person mailing paper or fee)

**CLAIM OF PRIORITY UNDER 37 C.F.R. § 1.78(a)(2)**

Mail Stop PCT  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Sir:

Pursuant to 35 U.S.C. 119, this application is entitled to a claim of priority to Japan  
Application No. 2003-130661 filed 08 May 2003.

Respectfully submitted,

FROMMER LAWRENCE & HAUG LLP  
Attorneys for Applicant

By: [Signature]  
William S. Frommer  
Reg. No. 25,506  
Tel. (212) 588-0800

Rec'd PCT/PTC 07 JAN 2005

2. 4. 2004

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

REC'D 21 MAY 2004

WIPO PCT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日                      2 0 0 3 年    5 月    8 日  
Date of Application:

出 願 番 号                      特 願 2 0 0 3 - 1 3 0 6 6 1  
Application Number:  
[ST. 10/C]:                      [ J P 2 0 0 3 - 1 3 0 6 6 1 ]

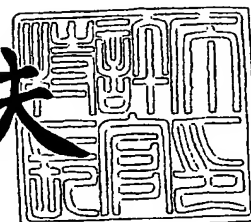
出      願      人                      ソニー株式会社  
Applicant(s):

PRIORITY DOCUMENT  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH  
RULE 17.1(a) OR (b)

2 0 0 4 年    2 月 2 7 日

特許庁長官  
Commissioner,  
Japan Patent Office

今 井 康 夫



BEST AVAILABLE COPY

出証番号    出証特 2 0 0 4 - 3 0 1 4 7 1 3

【書類名】 特許願  
【整理番号】 0390381507  
【提出日】 平成15年 5月 8日  
【あて先】 特許庁長官 太田 信一郎 殿  
【国際特許分類】 H04N 5/93  
G11B 20/10  
G11B 20/12

## 【発明者】

【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社  
内

【氏名】 加藤 元樹

## 【特許出願人】

【識別番号】 000002185

【氏名又は名称】 ソニー株式会社

## 【代理人】

【識別番号】 100067736

## 【弁理士】

【氏名又は名称】 小池 晃

## 【選任した代理人】

【識別番号】 100086335

## 【弁理士】

【氏名又は名称】 田村 榮一

## 【選任した代理人】

【識別番号】 100096677

## 【弁理士】

【氏名又は名称】 伊賀 誠司

## 【手数料の表示】

【予納台帳番号】 019530

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9707387

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 情報処理装置、情報処理方法、プログラム、及び記録媒体

【特許請求の範囲】

【請求項1】 トランспортパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、第1の多重化ストリーム最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続された多重化ストリームを復号する情報処理装置において、

上記多重化ストリームの上記アライバルタイムスタンプに従って上記ソースパケットを出力する出力手段と、

上記ソースパケットのうちビデオデータをバッファリングするビデオバッファと、

上記ソースパケットのうちオーディオデータをバッファリングするオーディオバッファと、

上記ビデオバッファにバッファリングされたビデオデータを復号するビデオ復号手段と、

上記オーディオバッファにバッファリングされたオーディオデータを復号するオーディオ復号手段とを有し、

上記オーディオバッファは、上記第2のピクチャを上記ビデオバッファに入力するために要する時間分のオーディオデータをバッファリング可能な容量を有することを特徴とする情報処理装置。

【請求項2】 上記オーディオバッファに必要な容量を $EBn\_max$  (bits) とし、上記第2のピクチャのビット量を $I\_max$  (bits) とし、上記ビデオバッファへの入力ビットレートを $Rv$  (bps) とし、オーディオデータのビットレートを $Ra$  (bps) としたとき、

$$EBn\_max = (I\_max / Rv) \times Ra$$

を満たすことを特徴とする請求項1記載の情報処理装置。

【請求項3】 上記第2のピクチャは、フレーム内符号化画像であることを特徴とする請求項1記載の情報処理装置。

【請求項4】 上記オーディオバッファは、少なくとも100ミリ秒分のオーディオデータをバッファリング可能な容量を有する

ことを特徴とする請求項1記載の情報処理装置。

【請求項5】 上記第1の多重化ストリームの時間軸における上記第1のピクチャの表示終了時刻と上記第2の多重化ストリームの時間軸における上記第2のピクチャの表示開始時刻との時間差をSTC\_deltaとし、上記第1の多重化ストリームの最後のソースパケットの最後のバイトが上記出力手段から出力される該第1の多重化ストリームの時間軸上の値STC1<sup>1</sup> endを上記時間差STC\_deltaにより上記第2の多重化ストリームの時間軸上の値に換算した値をSTC2<sup>1</sup> end (= STC1<sup>1</sup> end - STC\_delta) とし、上記第2の多重化ストリームの最初のパケットの最初のバイトが上記出力手段から出力される該第2の多重化ストリームの時間軸上の値をSTC2<sup>2</sup> startとしたとき、上記多重化ストリームは、  
$$STC2^2 \text{ start} > STC2^1 \text{ end}$$
を満たす

ことを特徴とする請求項1記載の情報処理装置。

【請求項6】 上記第1の多重化ストリームの時間軸における上記第1のピクチャの表示終了時刻と上記第2の多重化ストリームの時間軸における上記第2のピクチャの表示開始時刻との時間差をSTC\_deltaとし、上記第1の多重化ストリームの最後のソースパケットの最後のバイトが上記出力手段から出力される該第1の多重化ストリームの時間軸上の値STC1<sup>1</sup> endを上記時間差STC\_deltaにより上記第2の多重化ストリームの時間軸上の値に換算した値をSTC2<sup>1</sup> end (= STC1<sup>1</sup> end - STC\_delta) とし、上記第2の多重化ストリームの最初のパケットの最初のバイトが上記出力手段から出力される該第2の多重化ストリームの時間軸上の値をSTC2<sup>2</sup> startとし、上記第1の多重化ストリームの最後のパケットが上記出力手段から出力された後、所定時間delta1の経過後に上記第2の多重化ストリームの最初のパケットを上記出力手段から出力するものとしたとき、上記多重化ストリームは、  
$$STC2^2 \text{ start} > STC2^1 \text{ end} + \text{delta1}$$
を満たす

ことを特徴とする請求項1記載の情報処理装置。

【請求項7】 上記第1及び第2の多重化ストリームにおけるオーディオデータは、マルチチャンネルのオーディオデータからなる

ことを特徴とする請求項1記載の情報処理装置。

【請求項8】 トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、第1の多重化ストリームの最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続された多重化ストリームを復号する情報処理方法において、

上記多重化ストリームの上記アライバルタイムスタンプに従って上記ソースパケットを出力する出力工程と、

上記ソースパケットのうちビデオデータをビデオバッファにバッファリングし、オーディオデータをオーディオバッファにバッファリングするバッファリング工程と、

上記ビデオバッファ及びオーディオバッファにバッファリングされた夫々ビデオデータ及びオーディオデータを復号する復号工程とを有し、

上記バッファリング工程では、上記第2のピクチャを上記ビデオバッファに入力するために要する時間分のオーディオデータを、上記第2のピクチャを上記ビデオバッファにバッファリングする前に上記オーディオバッファにバッファリングする

ことを特徴とする情報処理方法。

【請求項9】 上記オーディオバッファの必要ビット量を $EBn\_max$  (bits) とし、上記第2のピクチャのビット量を $I\_max$  (bits) とし、上記ビデオバッファへの入力ビットレートを $Rv$  (bps)、オーディオデータのビットレートを $Ra$  (bps) としたとき、

$$EBn\_max = (I\_max / Rv) \times Ra$$

を満たすことを特徴とする請求項8記載の情報処理方法。

【請求項10】 トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、第1の多重化ストリー

ムの最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続された多重化ストリームを復号する処理をコンピュータに実行させるためのプログラムであって、

上記多重化ストリームの上記アライバルタイムスタンプに従って上記ソースパケットを出力する出力工程と、

上記ソースパケットのうちビデオデータをビデオバッファにバッファリングし、オーディオデータをオーディオバッファにバッファリングするバッファリング工程と、

上記ビデオバッファ及びオーディオバッファにバッファリングされた夫々ビデオデータ及びオーディオデータを復号する復号工程とを有し、

上記バッファリング工程では、上記第2のピクチャを上記ビデオバッファに入力するために要する時間分のオーディオデータを、上記第2のピクチャを上記ビデオバッファにバッファリングする前に上記オーディオバッファにバッファリングする

ことを特徴とするプログラム。

【請求項11】 上記オーディオバッファの必要ビット量を $EBn\_max$  (bits) とし、上記第2のピクチャのビット量を $I\_max$  (bits) とし、上記ビデオバッファへの入力ビットレートを $Rv$  (bps)、オーディオデータのビットレートを $Ra$  (bps) としたとき、

$$EBn\_max = (I\_max / Rv) \times Ra$$

を満たすことを特徴とする請求項10記載のプログラム。

【請求項12】 トランSPORTパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、第1の多重化ストリームの最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続された多重化ストリームを復号する処理をコンピュータに実行させるためのプログラムが記録されたコンピュータ読み取り可能な記録媒体であって、

上記多重化ストリームの上記アライバルタイムスタンプに従って上記ソースパケットを出力する出力工程と、



上記ソースパケットのうちビデオデータをビデオバッファにバッファリングし、オーディオデータをオーディオバッファにバッファリングするバッファリング工程と、

上記ビデオバッファ及びオーディオバッファにバッファリングされた夫々ビデオデータ及びオーディオデータを復号する復号工程とを有し、

上記バッファリング工程では、上記第2のピクチャを上記ビデオバッファに入力するために要する時間分のオーディオデータを、上記第2のピクチャを上記ビデオバッファにバッファリングする前に上記オーディオバッファにバッファリングする

ことを特徴とするプログラムが記録された記録媒体。

【請求項13】 上記オーディオバッファの必要ビット量を $EBn\_max$  (bits)とし、上記第2のピクチャのビット量を $I\_max$  (bits)とし、上記ビデオバッファへの入力ビットレートを $Rv$ (bps)、オーディオデータのビットレートを $Ra$  (bps)としたとき、

$$EBn\_max = (I\_max / Rv) \times Ra$$

を満たすことを特徴とする請求項12記載の記録媒体。

【請求項14】 トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなる多重化ストリームが記録された記録媒体であって、

上記多重化ストリームは、第1の多重化ストリームの最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続され、上記第1及び第2の多重化ストリームが夫々のアライバルタイムスタンプに基づいてデコーダに入力可能であって、且つ、上記第2のピクチャを該デコーダに入力するために要する時間分のオーディオデータを該第2のピクチャが該デコーダに入力開始される前までに入力終了可能なように多重化された多重化ストリームが記録された記録媒体。

【請求項15】 上記第2のピクチャを上記デコーダに入力するために要する時間分のオーディオデータは、上記第2のピクチャのビット量を $I\_max$  (bits)とし、上記デコーダのビデオバッファへの入力ビットレートを $Rv$ (bps)、オーディオ

ィオデータのビットレートを $R_a$  (bps) としたとき、 $(I_{\max}/R_v) \times R_a$ であることを特徴とする請求項14記載の記録媒体。

【請求項16】 上記第2のピクチャは、フレーム内符号化画像であることを特徴とする請求項14記載の記録媒体。

【請求項17】 トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、該アライバルタイムスタンプに基づきデコーダにより読み出されてデコードされる多重化ストリームを生成する情報処理装置において、

第1のピクチャで表示終了する第1のビデオ符号化ストリームを生成し、この第1のピクチャに続けて表示される第2のピクチャから表示開始する第2のビデオ符号化ストリームを生成するビデオ符号化手段と、

上記第1のビデオ符号化ストリームとこの上記第1のビデオ符号化ストリームに同期したオーディオ符号化ストリームとを多重化して第1の多重化ストリームを生成し、上記第2のビデオ符号化ストリームとこの第2のビデオ符号化ストリームに同期したオーディオ符号化ストリームとを多重化して第2の多重化ストリームを生成し、上記第1の多重化ストリームの最後のピクチャである上記第1のピクチャに上記第2の多重化ストリームの最初のピクチャである上記第2のピクチャが連続して再生されるよう接続された多重化ストリームを生成する多重化手段とを有し、

上記多重化手段は、上記第2のピクチャを上記デコーダに入力するために要する時間分のオーディオデータを上記第2のピクチャが該デコーダに入力開始される前までに入力終了可能なように多重化する

ことを特徴とする情報処理装置。

【請求項18】 上記第2のピクチャを上記デコーダに入力するために要する時間分のオーディオデータは、上記第2のピクチャのビット量を $I_{\max}$  (bits) とし、上記デコーダのビデオバッファへの入力ビットレートを $R_v$  (bps)、オーディオデータのビットレートを $R_a$  (bps) としたとき、 $(I_{\max}/R_v) \times R_a$ であることを特徴とする請求項17記載の情報処理装置。

【請求項19】 上記第2のピクチャは、フレーム内符号化画像であることを

特徴とする請求項17記載の情報処理装置。

【請求項20】    トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、該アライバルタイムスタンプに基づきデコーダにより読み出されてデコードされる多重化ストリームを生成する情報処理方法において、

第1のピクチャで表示終了する第1のビデオ符号化ストリームを生成し、この第1のピクチャに続けて表示される第2のピクチャから表示開始する第2のビデオ符号化ストリームを生成するビデオ符号化工程と、

上記第1のビデオ符号化ストリームとこの上記第1のビデオ符号化ストリームに同期したオーディオ符号化ストリームとを多重化して第1の多重化ストリームを生成し、上記第2のビデオ符号化ストリームとこの第2のビデオ符号化ストリームに同期したオーディオ符号化ストリームとを多重化して第2の多重化ストリームを生成し、上記第1の多重化ストリームの最後のピクチャである上記第1のピクチャに上記第2の多重化ストリームの最初のピクチャである上記第2のピクチャが連続して再生されるよう接続された多重化ストリームを生成する多重化工程とを有し、

上記多重化工程では、上記第2のピクチャを上記デコーダに入力するために要する時間分のオーディオデータを上記第2のピクチャが該デコーダに入力開始される前までに入力終了可能なように多重化する

ことを特徴とする情報処理方法。

#### 【発明の詳細な説明】

##### 【0001】

##### 【発明の属する技術分野】

本発明は、ビデオストリームとオーディオストリームとの多重化ストリームをビデオフレーム精度で編集して、編集点を連続的（シームレス）に再生するための情報処理装置、その方法、プログラム及び記録媒体、並びにシームレス再生に最適な形式で多重化ストリームを生成する情報処理装置及びその多重化されたストリームデータが記録された記録媒体に関する。

##### 【0002】

**【従来の技術】**

ビデオストリームとオーディオストリームの多重化ストリームをビデオフレーム精度で編集して、編集点をシームレスに再生するための方法が下記特許文献1乃至3に記載されている。

**【0003】**

図12は、従来のDVR-STDモデル (DVR MPEG2 transport stream player model) (以下、プレーヤという。) 101を示すブロック図である。DVR-STDは、シームレス接続された2つのPlayItemによって参照されるAVストリームの生成および検証の際におけるデコード処理をモデル化するための概念モデルである。

**【0004】**

図12に示すように、プレーヤ101においては、読出部 (DVRdrive) 111からビットレートRUDで読み出されたTS (トランスポートストリーム) ファイルは、リードバッファ112にバッファリングされ、このリードバッファ112からソースパケットがソースデパケッタイザ部 (1 source depacketizer) 113へ最大ビットレートRMAXで読み出される。

**【0005】**

パルス発振器 (27MHz X-tal) 114は、27MHzのパルスを発生する。アライバルクロックカウンタ (Arrival time clock counter) 115は、この27MHzの周波数のパルスをカウントするバイナリカウンタであり、ソースデパケッタイザ部113に、時刻 $t(i)$ におけるArrival time clock counterのカウント値Arrival\_time\_clock(i)を供給する。

**【0006】**

1つのソースパケットは、1つのトランスポートパケットとそのarrival\_time\_stampを持つ。現在のソースパケットのarrival\_time\_stampがarrival\_time\_clock(i)のLSB (least significant Bit: 最下位ビット) 30ビットの値と等しい時、ソースデパケッタイザ部113からそのトランスポートパケットが出力される。TS\_recording\_rateは、トランスポートストリーム (以下、TSという。) のビットレートである。また、図12に示すn、TBn、MBn、EBn、TBSys、BSys、Rxn、Rbxn、Rxsys、Dn、Dsys、On及びPn(k)の表記方法は、ISO/IEC13818-1 (M

PEG2 Systems規格)のT-STD (ISO/IEC 13818-1で規定されるtransport stream system target decoder) に定義されているものと同じである。

#### 【0007】

次に、このような従来のプレーヤ101におけるデコーディングプロセスについて説明する。まず、1つのDVR MPEG2 TSを再生している時のデコーディングプロセスを説明する。単一のDVR MPEG2TSを再生している間は、出力部110からトランスポート packets を出力してデコーダ120であるDVR-STDのTB1、TBn又はTBsysのバッファへ入力するタイミングは、ソース packets のarrival\_time\_stampにより決定される。TB1、MB1、EB1、TBn、Bn、TBsys及びTBsysのバッファリミング動作の規定は、ISO/IEC 13818-1に規定されているT-STDと同じである。復号動作と表示動作の規定もまた、ISO/IEC 13818-1に規定されているT-STDと同じである。

#### 【0008】

次に、シームレス接続されたPlayItemを再生している間のデコーディングプロセスについて説明する。ここでは、シームレス接続されたPlayItemによって参照される先行するストリームTS1と、現在のストリームTS2との再生について説明をする。

#### 【0009】

あるAVストリーム (TS1) からそれにシームレスに接続された次のAVストリーム (TS2) へと移る間には、TS2のアライバルタイムベースの時間軸は、TS1のアライバルタイムベースの時間軸と同じ時間軸でない。また、TS2のシステムタイムベースの時間軸は、TS1のシステムタイムベースの時間軸と同じ時間軸でない。ビデオの表示は、シームレスに連続していることが要求される。オーディオのプレゼンテーションユニットの表示時間にはオーバーラップがあってもよい。

#### 【0010】

次に、ソースデパケッタ部から読み出されるトランスポート packets のDVR-STDへの入力タイミングについて説明する。

#### 【0011】

- (1) TS1の最後のビデオ packets がDVR-STDのTB1に入力終了する時刻T1ま

での時間

時刻  $T_1$  までの時間は、DVR-STDのTB1、TBn又はTBsysのバッファへの入力タイミングは、TS1のソースパケットのarrival\_time\_stampによって決定される。

【0012】

(2) 時刻  $T_1$  からTS1の残りのパケットの最後のバイトがに入力終了する時刻  $T_2$  まで

TS1の残りのパケットは、TS\_recording\_rate (TS1) のビットレート (TS1の最大ビットレート) でDVR-STDのTBn又はTBsysのバッファへ入力されなければならない。ここで、TS\_recording\_rate (TS1) は、Clip1に対応するClipInfo()において定義されるTS\_recording\_rateの値である。TS1の最後のバイトがバッファへ入力する時刻は、時刻  $T_2$  である。従って、時刻  $T_1$  から  $T_2$  までの区間では、ソースパケットのarrival\_time\_stampは無視される。

【0013】

$N_1$  をTS1の最後のビデオパケットに続くTS1のトランスポートパケットのバイト数とすると、時刻  $T_1$  から  $T_2$  までの時間  $T_2 - T_1$  は、 $N_1$  バイトがTS\_recording\_rate (TS1) のビットレートで入力終了するために必要な時間であり、下記式 (1) のように計算される。

【0014】

【数1】

$$T_{2-1} = T_2 - T_1 = N_1 / \text{TS\_recording\_rate (TS1)} \cdots (1)$$

【0015】

時刻  $T_1$  から  $T_2$  までの間は、図12に示すRxnとRxsysの値は共にTS\_recording\_rate (TS1) の値に変化する。このルール以外のバッファリング動作は、T-STDと同じである。

【0016】

オーディオデコーダは、時刻  $T_1$  から  $T_2$  までの区間の入力データ进行处理することができるように、即ち、時刻  $T_1$  から  $T_2$  までの間は、図12に示すRxnとR

xsysの値が共にTS\_recording\_rate (TS1) の値に変化するため、T-STDで定義されるバッファ量に加えて付加的なバッファ量 (約1秒分のデータ量) が必要である。

### 【0017】

(3) 時刻T2以後

T2の時刻において、アライバルタイムクロックカウンタ115は、TS2の最初のソースパケットのarrival\_time\_stampの値にリセットされる。DVR-STDのTB1、TBn又はTBsysのバッファへの入力タイミングは、TS2のソースパケットのarrival\_time\_stampによって決定される。RxnとRxsysは共に、T-STDにおいて定義されている値に変化する。

### 【0018】

次に、ビデオのプレゼンテーションタイミングについて説明する。ビデオプレゼンテーションユニットの表示は、接続点 (コネクションポイント) を通してギャップなしに連続でなければならない。

ここで、

STC (System Time Clock) 1 : TS1のシステムタイムベースの時間軸

STC2 : TS2のシステムタイムベースの時間軸 (正確には、STC2は、TS2の最初のPC

R (Program Clock Reference) がT-STDに入力した時刻から開始する。)

とする。

### 【0019】

STC1とSTC2との間のオフセット値は、次のように決定される。

PTS1\_end : TS1の最後のビデオプレゼンテーションユニットに対応するSTC1上のPTS

PTS2\_start : TS2の最初のビデオプレゼンテーションユニットに対応するSTC2上のPTS

T\_pp : 最後のビデオプレゼンテーションユニットの表示期間

とすると、

2つのシステムタイムベースの間のオフセット値STC\_deltaは、下記式 (2) のように計算される。

【0020】

【数2】

$$STC\_delta = PTS1_{end} + T_{pp} - PTS2_{start} \quad \dots (2)$$

【0021】

次に、オーディオのプレゼンテーションタイミングについて説明する。TS1とTS2との接続点において、オーディオプレゼンテーションユニットの表示タイミングのオーバーラップがあってもよく、それは0から2オーディオフレーム未満である。プレーヤ101には、どちらのオーディオサンプルを選択するかということと、オーディオプレゼンテーションユニットの表示を接続点の後の補正されたタイムベースに再同期する制御が要求される。

【0022】

TS1からそれにシームレスに接続された次のTS2へと移る時、DVR-STDのシステムタイムクロックの制御について、プレーヤ101が行う処理を説明する。TS1の最後のオーディオプレゼンテーションユニットが表示される時刻T5において、システムタイムクロックは、時刻T2からT5の間にオーバーラップしていてもよい。この区間では、DVR-STDは、システムタイムクロックを古いタイムベースの値(STC1)から新しいタイムベースの値(STC2)の値に切り替える。STC2の値は、下記式(3)のように計算できる。

【0023】

【数3】

$$STC2 = STC1 - STC\_delta \quad \dots (3)$$

【0024】

TS1から、これにシームレスに接続された次のTS2へと移る時、TS1及びTS2が満たさなければいけない符号化条件を説明する。

【0025】

$STC1^{1}_{video\_end}$ : TS1の最後のビデオパケットの最後のバイトがDVR-STDのTB1へ到着する時のシステムタイムベースSTC1上のSTCの値



STC2<sup>2</sup><sub>video\_start</sub>: TS2の最初のビデオパケットの最初のバイトがDVR-STDのTB1へ到着する時のシステムタイムベースSTC2上のSTCの値  
 STC2<sup>1</sup><sub>video\_end</sub>: STC1<sup>1</sup><sub>video\_end</sub>の値をシステムタイムベースSTC2上の値に換算した値  
 とした場合、STC2<sup>1</sup><sub>video\_end</sub>は、下記式(4)のようにして計算される。

【0026】

【数4】

$$STC2^1_{video\_end} = STC1^1_{video\_end} - STC\_delta \quad \dots (4)$$

【0027】

ここで、デコーダ120がDVR-STDに従うために、次の2つの条件を満たすことが要求される。

【0028】

(条件1)

TS2の最初のビデオパケットのTB1への到着タイミングは、次に示す不等式(5)を満たさなければならない。

【0029】

【数5】

$$STC2^2_{video\_start} > STC2^1_{video\_end} + T_{2-1} \quad \dots (5)$$

【0030】

この上記不等式(5)を満たすように、Clip1及び／又はClip2の部分的なストリームを再エンコード及び／又は再多重化することが必要になる。

【0031】

(条件2)

STC1とSTC2とを同じ時間軸上に換算したシステムタイムベースの時間軸上にお

いて、TS1からのビデオパケットの入力とそれに続くTS2からのビデオパケットの入力は、ビデオバッファをオーバーフロー及びアンダーフローさせてはならない。

### 【0032】

#### 【特許文献1】

特開2000-175152号公報

#### 【特許文献2】

特開2001-544118号公報

#### 【特許文献3】

特開2002-158974号公報

### 【0033】

#### 【発明が解決しようとする課題】

しかしながら、上述したように、DVR-STDモデルを使用した従来のプレーヤ101においては、時刻T1からT2までの区間の入力データを処理することができると、即ち、時刻T1からT2までの間は、TS1の残りのパケットは、TS\_recording\_rate (TS1) のビットレート (TS1の最大ビットレート) でDVR-STDのTBn又はTBsysのバッファへ入力されていたため、T-STDで定義されるバッファに加えて、約1秒分のデータをバッファリング可能な容量の付加的なバッファが必要である。

### 【0034】

このバッファ容量の大きさは、次の要因に基づく。すなわち、MPEG2TSの中心において、あるバイト位置にあるビデオデータに同期再生されるオーディオデータが、所定範囲内の多重化位相差を離れて存在することができ、この多重化位相差の最大値が1秒分のデータ量に相当する。従って、上記式(1)のN1の最大値は、最大1秒分のオーディオデータに相当する。時刻T1からT2までの間は、ソースパケットのarrival\_time\_stampを無視して、TSの最大ビットレートでN1のデータ量のソースパケットが、オーディオバッファに入力されるので、このデータ量をバッファリングするために、T-STDで定義されるバッファ量に加えて付加的なバッファ量 (約1秒分のデータ量) が必要になっていた。

## 【0035】

この付加的なバッファの大きさを具体的に計算すると次のようになる。即ち、例えば640kbpsのドルビーAC3方式により符号化されたオーディオストリームの場合、1秒分のオーディオデータは、640kbits=80kBytesとなり、80kBytesの付加的なバッファが必要となる。

## 【0036】

また、Linear PCM方式により符号化されたオーディオストリームであって、24bitSample、96kHz sampling frequency、8 channelsの場合、1秒分のオーディオデータは、 $24\text{bitSample} \times 96000\text{samples/sec} \times 8\text{channels} = \text{約}18\text{Mbits}$ となり、約3Mbytesの付加的なバッファが必要となり、このようなマルチチャンネルのオーディオデータを扱う場合、この付加的なバッファが極めて大きくなってしまいうという問題点がある。

## 【0037】

本発明は、このような従来の実情に鑑みて提案されたものであり、オーディオストリームとビデオストリームとが多重化された2つの多重化ストリームをシームレスに連続して復号するために、最適な容量のオーディオバッファとした情報処理装置、その方法、プログラム及び記録媒体、並びにこのようなオーディオバッファ容量に対応した多重化ストリームを生成する情報処理装置及その方法、並びにその多重化ストリームが記録された記録媒体を提供することを目的とする。

## 【0038】

## 【課題を解決するための手段】

上述した目的を達成するために、本発明に係る情報処理装置は、トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、第1の多重化ストリームの最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続された多重化ストリームを復号する情報処理装置において、上記多重化ストリームの上記アライバルタイムスタンプに従って上記ソースパケットを出力する出力手段と、上記ソースパケットのうちビデオデータをバッファリングするビデオバッファと、上記ソースパケットのうちオーディオデー

タをバッファリングするオーディオバッファと、上記ビデオバッファにバッファリングされたビデオデータを復号するビデオ復号手段と、上記オーディオバッファにバッファリングされたオーディオデータを復号するオーディオ復号手段とを有し、上記オーディオバッファは、上記第2のピクチャを上記ビデオバッファに入力するために要する時間分のオーディオデータをバッファリング可能な容量を有することを特徴とする。

#### 【0039】

本発明においては、オーディオバッファのサイズを、上記第2のピクチャを上記ビデオバッファに入力するために要する時間分のオーディオデータをバッファリング可能な容量とすると共に、上記第1のピクチャが上記ビデオバッファに入力終了する時刻から第1の多重化ストリームの最後の packets を入力終了する時刻までの間においても、多重化ストリームのソース packets のアライバルタイムスタンプ (arrival\_time\_stamp) に従ってソース packets をバッファに入力するようにしたので、従来ソース packets の arrival\_time\_stamp を無視して、トランスポートストリーム (TS) の最大ビットレートで入力するために必要とされていた1秒分の付加的なバッファを不要とすると共に、第1の多重化ストリームの最後のトランスポート packets を入力した後に、第2の多重化ストリームの最初にデコードされるピクチャをそのデコードタイミングまでにビデオバッファへ入力することができる。

#### 【0040】

また、上記オーディオバッファの必要な容量を  $EBn\_max$  (bits) とし、上記第2のピクチャのビット量を  $I\_max$  (bits) とし、上記ビデオバッファへの入力ビットレートを  $Rv$  (bps) とし、オーディオデータのビットレートを  $Ra$  (bps) としたとき、 $EBn\_max = (I\_max / Rv) \times Ra$  を満たすものとしてことができ、第2のピクチャとなる例えば  $I$  ピクチャのビット量の最大値を  $I\_max$  とした場合、オーディオバッファの容量は、最大で  $EBn\_max = (I\_max / Rv) \times Ra$  とすることができる。

#### 【0041】

更に、上記オーディオバッファは、100ミリ秒分のオーディオデータをバッファリング可能な容量を有することが好ましく、これにより、例えばMPEG2の  $I$

ピクチャのデータ量は、通常1秒間で伝送されるデータ量の10%以下の大きさであるため、オーディオバッファをこれと同じ大きさの容量としておき、その大きさ分のオーディオデータを先送りしておくことで、Iピクチャをそのデコードタイミングまでにビデオバッファへ入力することができ、ビデオデータのエンコード制限が少なくなる。即ち、オーディオバッファを上記容量とすることで、オーディオデータをそれが再生されるタイミングよりも100ミリ秒だけ早く入力終了するように多重化ストリームを多重化することができる。

#### 【0042】

更にまた、上記第1の多重化ストリームの時間軸における上記第1のピクチャの表示終了時刻と上記第2の多重化ストリームの時間軸における上記第2のピクチャの表示開始時刻との時間差をSTC\_deltaとし、上記第1の多重化ストリームの最後のソースパケットの最後のバイトが上記出力手段から出力される該第1の多重化ストリームの時間軸上の値 $STC1^{1\text{end}}$ を上記時間差STC\_deltaにより上記第2の多重化ストリームの時間軸上の値に換算した値を $STC2^{1\text{end}}$  ( $=STC1^{1\text{end}} - STC\_delta$ ) とし、上記第2の多重化ストリームの最初のパケットの最初のバイトが上記出力手段から出力される該第2の多重化ストリームの時間軸上の値を $STC2^{2\text{start}}$ としたとき、上記多重化ストリームは、 $STC2^{2\text{start}} > STC2^{1\text{end}}$ を満たすものとするにより、DVR-STDに従ったものとすることができる。

#### 【0043】

更にまた、上記第1の多重化ストリームの最後のパケットが上記出力手段から出力された後、所定時間delta1の経過後、上記第2の多重化ストリームの最初のパケットを上記出力手段から出力するものとし、 $STC2^{2\text{start}} > STC2^{1\text{end}} + delta1$ を満たすものとしてもよく、これにより、第2の多重化ストリームの最初のパケットの入力タイミングに柔軟性を持たせ、第2の多重化ストリームの符号化を容易にすることができる。

#### 【0044】

本発明に係る情報処理方法は、トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、第1の多

重化ストリームの最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続された多重化ストリームを復号する情報処理方法において、上記多重化ストリームの上記アライバルタイムスタンプに従って上記ソース packets を出力する出力工程と、上記ソース packets のうちビデオデータをビデオバッファにバッファリングし、オーディオデータをオーディオバッファにバッファリングするバッファリング工程と、上記ビデオバッファ及びオーディオバッファにバッファリングされたビデオデータ及びオーディオデータを復号する復号工程とを有し、上記バッファリング工程では、上記第2のピクチャを上記ビデオバッファに入力するために要する時間分のオーディオデータを、上記第2のピクチャを上記ビデオバッファにバッファリングする前に上記オーディオバッファにバッファリングすることを特徴とする。

#### 【0045】

また、本発明に係るプログラムは、上述した情報処理をコンピュータに実行させるものであり、本発明に係る記録媒体は、そのようなプログラムが記録されたコンピュータ読み取り可能なものである。

#### 【0046】

本発明に係る他の記録媒体は、トランスポート packets とそのアライバルタイムスタンプとを有するソース packets を単位とするデータ列からなる多重化ストリームが記録された記録媒体であって、上記多重化ストリームは、第1の多重化ストリームの最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続され、上記第1及び第2の多重化ストリームが夫々のアライバルタイムスタンプに基づいてデコーダに入力可能であって、且つ、上記第2のピクチャをデコーダに入力するために要する時間分のオーディオデータを該第2のピクチャが該デコーダに入力開始される前までに入力終了可能なように多重化された多重化ストリームが記録されたことを特徴とする。

#### 【0047】

本発明においては、第2のピクチャを上記デコーダに入力するために要する時

間分のオーディオデータを該第2のピクチャが該デコーダに入力開始される前までに入力終了可能なように多重化ストリームが多重化されているため、このような多重化ストリームを、第2のピクチャをビデオバッファに入力するために要する時間分のオーディオデータをバッファリング可能な容量のオーディオバッファを有するデコーダにより復号すれば、第1の多重化ストリームの最後のトランスポートパケットを入力した後に、第2の多重化ストリームの最初にデコードされるピクチャをそのデコードタイミングまでにビデオバッファへ入力することができる。

#### 【0048】

本発明に係る他の情報処理装置は、トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、該アライバルタイムスタンプに基づきデコーダにより読み出されてデコードされる多重化ストリームを生成する情報処理装置において、第1のピクチャで表示終了する第1のビデオ符号化ストリームを生成し、この第1のピクチャに続けて表示される第2のピクチャから表示開始する第2のビデオ符号化ストリームを生成するビデオ符号化手段と、上記第1のビデオ符号化ストリームとこの上記第1のビデオ符号化ストリームに同期したオーディオ符号化ストリームとを多重化して第1の多重化ストリームを生成し、上記第2のビデオ符号化ストリームとこの第2のビデオ符号化ストリームに同期したオーディオ符号化ストリームとを多重化して第2の多重化ストリームを生成し、上記第1の多重化ストリームの最後のピクチャである上記第1のピクチャに上記第2の多重化ストリームの最初のピクチャである上記第2のピクチャが連続して再生されるよう接続された多重化ストリームを生成する多重化手段とを有し、上記多重化手段は、上記第2のピクチャを上記デコーダに入力するために要する時間分のオーディオデータを上記第2のピクチャが該デコーダに入力開始される前までに入力終了可能なように多重化することを特徴とする。

#### 【0049】

本発明においては、上記第2のピクチャを上記デコーダに入力するために要する時間分の、例えば100ミリ秒分のオーディオデータを上記第2のピクチャが

該デコーダに入力開始される前までに入力終了可能なように多重化するため、デコーダにおいて、オーディオデータをオーディオバッファに先送りして、Iピクチャ等の第2のピクチャを、そのデコードタイミングまでに伝送する時間を十分に確保することができ、多重化ストリームの符号化が容易になる。

#### 【0050】

本発明に係る他の情報処理方法は、トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、該アライバルタイムスタンプに基づきデコーダにより読み出されてデコードされる多重化ストリームを生成する情報処理方法において、第1のピクチャで表示終了する第1のビデオ符号化ストリームを生成し、この第1のピクチャに続けて表示される第2のピクチャから表示開始する第2のビデオ符号化ストリームを生成するビデオ符号化工程と、上記第1のビデオ符号化ストリームとこの上記第1のビデオ符号化ストリームに同期したオーディオ符号化ストリームとを多重化して第1の多重化ストリームを生成し、上記第2のビデオ符号化ストリームとこの第2のビデオ符号化ストリームに同期したオーディオ符号化ストリームとを多重化して第2の多重化ストリームを生成し、上記第1の多重化ストリームの最後のピクチャである上記第1のピクチャに上記第2の多重化ストリームの最初のピクチャである上記第2のピクチャが連続して再生されるよう接続された多重化ストリームを生成する多重化工程とを有し、上記多重化工程では、上記第2のピクチャを上記デコーダに入力するために要する時間分のオーディオデータを上記第2のピクチャが該デコーダに入力開始される前までに入力終了可能なように多重化することを特徴とする。

#### 【0051】

##### 【発明の実施の形態】

以下、本発明を適用した具体的な実施の形態について、図面を参照しながら詳細に説明する。この実施の形態は、本発明を、ビデオストリームとオーディオストリームとが多重化された2つのAVストリームをシームレスで連続再生する情報処理装置に適用したものである。そして、本実施の形態においては、DVR-STD (Digital Video Recording-System Target Decoder) において、シームレスに接続



された2つのAVストリームを再生する際に最適な容量としたオーディオバッファを提案するものである。

#### 【0052】

はじめに、以下の説明において使用する用語について説明する。Clipは、ビデオストリームとオーディオストリームとの多重化ストリームを示す。また、Play Listは、Clipの中の再生区間の集まりを示す。あるClipの中の一つの再生区間は、PlayItemと呼ばれ、それは、時間軸上のIN点とOUT点のペアで表される。それゆえ、PlayListは、PlayItemの集まりである。

#### 【0053】

PlayItem間をシームレスに再生するとは、再生装置（プレーヤ）が、デコーダの再生出力にポーズやギャップを起こさせることなく、ディスクに記録されたオーディオビデオデータを表示（再生）することができることである。

#### 【0054】

シームレス接続されている2つのPlayItemsの構造について説明する。先行するPlayItemと現在のPlayItemとの接続が、シームレス表示できるように保証されているかどうかは、現在のPlayItemにおいて定義されているconnection\_conditionフィールドから判断することができる。そして、PlayItem間のシームレス接続は、Bridge-Clip (BridgeSequence) を使用する方法（以下、第1の方法とする。）と使用しない方法（以下、第2の方法とする。）がある。

#### 【0055】

まず、先行するPlayItem (previous PlayItem) と現在のPlayItem (current Play Item) とが、第1の方法であるBridgeSequenceを使用して接続されている場合のTS1及びTS2について説明する。図1は、第1の方法であるBridge-Clipを使用する場合の先行するPlayItemと現在のPlayItemとの関係を示す模式図である。この図1において、Bridge-Clipを使用する場合にプレーヤが読み出すストリームデータを影を付けて示す。ここで、DVR MPEG (Moving Picture Experts Group) 2トランスポートストリーム (TS) は、整数個のAligned unitから構成される。Aligned unitの大きさは、6144バイト(2048×3バイト)である。1つのAligned unitは、32個のソースパケットからなり、ソースパケットの第1バイト

目から始まる。

#### 【0056】

各ソースパケットは、192バイト長である。1つのソースパケットは、TP\_extra\_headerとトランスポートパケットとから成り、TP\_extra\_headerは、4バイト長であり、またトランスポートパケットは、188バイト長である。TP\_extra\_headerは、copy\_premission\_indicatorとarrival\_time\_stampとを有し、copy\_premission\_indicatorは、トランスポートパケットのペイロード (Payload) のコピー制限を表す整数、arrival\_time\_stamp (ATS) は、AVストリームの中で、対応するトランスポートパケットがデコードに到着する時刻を示すタイムスタンプである。AVストリームを構成する各ソースパケットのarrival\_time\_stampに基づいて作られる時間軸をアライバルタイムベースといい、そのクロックをATC (Arrival Time Clock) と呼ぶ。

#### 【0057】

図1に示すTS1 (第1の多重化ストリーム) は、Clip1 (Clip AVストリーム) の影を付けて示すストリームデータD1とBridge-ClipのSPN\_arrival\_time\_discontinuityより前の影を付けて示すストリームデータD2からなる。SPN\_arrival\_time\_discontinuityは、the Bridge-Clip AVストリームファイルの中でアライバルタイムベースの不連続点があるところのソースパケットのアドレスを示す。

#### 【0058】

そして、TS1に含まれるClip1の影を付けて示すストリームデータD1は、先行するPlayItemのIN\_time (図1においてIN\_time1で示す。) に対応するプレゼンテーションユニットを復号するために必要なストリームのアドレスから、SPN\_exit\_from\_previous\_Clipで参照されるソースパケットまでのストリームデータである。

#### 【0059】

また、TS1に含まれるBridge-ClipのSPN\_arrival\_time\_discontinuityより前の影を付けて示すストリームデータD2は、Bridge-Clipの最初のソースパケットから、SPN\_arrival\_time\_discontinuityで参照されるソースパケットの直前のソースパケットまでのストリームデータである。

## 【0060】

また、図1に示すTS2（第2の多重化ストリーム）は、Clip2（Clip AVストリーム）の影を付けて示すストリームデータD4とBridge-ClipのSPN\_arrival\_time\_discontinuity以後の影を付けて示すストリームデータD3からなる。

## 【0061】

そして、TS2に含まれるBridge-ClipのSPN\_arrival\_time\_discontinuity以後の影を付けて示すストリームデータD3は、SPN\_arrival\_time\_discontinuityで参照されるソースパケットから、Bridge-Clipの最後のソースパケットまでのストリームデータである。

## 【0062】

また、TS2に含まれるClip2の影を付けて示すストリームデータD4は、SPN\_enter\_to\_current\_Clipで参照されるソースパケットから、現在のPlayItemのOUT\_time（図1において、OUT\_time2で示す。）に対応するプレゼンテーションユニットを復号するために必要なストリームのアドレスまでのストリームデータである。

## 【0063】

次に、先行するPlayItemと現在のPlayItemが、第2の方法であるBridgeSequenceを使用しないで接続されている場合のTS1及びTS2について説明する。図2は、第2の方法であるBridge-Clipを使用しない場合の先行するPlayItemと現在のPlayItemの関係を示す模式図である。この図2において、プレーヤが読み出すストリームデータは、影を付けて示す。

## 【0064】

図2に示すTS1（第1の多重化ストリーム）は、Clip1（Clip AVストリーム）の影を付けて示すストリームデータD5からなる。TS1に含まれるClip1の影を付けて示すストリームデータD5は、先行するPlayItemのIN\_time（図2においてIN\_time1で示す）に対応するプレゼンテーションユニットを復号するために必要なストリームのアドレスから始まり、Clip1の最後のソースパケットまでのデータである。

## 【0065】

また、図2に示すTS2（第2の多重化ストリーム）は、Clip2（Clip AVストリーム）の影を付けて示すストリームデータD6からなる。TS2に含まれるClip2の影を付けて示すストリームデータD6は、Clip2の最初のソースパケットから始まり、現在のPlayItemのOUT\_time（図2においてOUT\_time2で示す）に対応するプレゼンテーションユニットを復号するために必要なストリームのアドレスまでのストリームデータである。

#### 【0066】

図1及び図2の両方において、TS1及びTS2は、ソースパケットの連続したストリームである。次に、TS1及びTS2のストリーム規定とそれらの間の接続条件について説明する。

#### 【0067】

TS1及びTS2は、ビデオストリームとオーディオストリームとが多重化されたものであるが、ここでは、まず、シームレス接続のための符号化制限におけるビデオビットストリームの制限について説明する。

#### 【0068】

図3は、ビデオストリームとしてのClip1とClip2とをシームレス接続する例をピクチャの表示順序（Presentation order）で示す模式図である。動画像プログラム的一部分をスキップ再生する際に、スキップ再生開始点であるアウト点ピクチャより時間的に前側のプログラムであるアウト点側プログラムと、スキップ再生到達点であるイン点ピクチャより時間に後側のプログラムであるイン点側プログラムとをシームレスに接続するには、復号装置においてビデオストリームの再エンコード処理を行う必要がある。

#### 【0069】

MPEG2規格に準じた画像群の単位であるGOP（group of pictures）には、他の画像からの予測符号化なしに画像が符号化された参照画像である少なくとも1つのI（Intra）ピクチャ（フレーム内符号化画像）と、表示順序に順方向の予測符号化を用いて画像が符号化された順方向予測符号化画像であるP（predictive）ピクチャと、順方向及び逆方向の予測符号化を用いて画像が符号化された双方向予測符号化画像であるB（bidirectionally）ピクチャとの3種類の符号化

画像が含まれている。図3においては、Clip1及びClip2の各数字は表示順序を示し、I、P、B、又はi、p、bは各ピクチャの種類を示す。例えば図3には、Clip1のB7とClip2のb4とを接続する例を示すが、この接続点においてビデオストリームをシームレスに表示できるためには、OUT\_time1 (Clip1のOUT\_time) の後とIN\_time2 (Clip2のIN\_time) の前に表示される不必要なピクチャは、接続点付近のClipの部分的なストリームを再エンコードするプロセスにより、除去されなければならない。

#### 【0070】

図3に示すようなビデオストリーム (Clip1及びClip2) をシームレス接続する場合に、上記第1の方法であるBridgeSequenceを使用してシームレス接続を実現する例を図4に示す。SPN\_arrival\_time\_discontinuityより前のBridge-Clipのビデオストリームは、図3に示すClip1のOUT\_time1に対応するピクチャまでの符号化ビデオストリームからなる。そして、そのビデオストリームは先行するClip1のビデオストリームに接続され、一つの連続でMPEG2規格に従ったエレメンタリーストリームとなるように再エンコードされている。同様に、SPN\_arrival\_time\_discontinuity以後のBridge-Clipのビデオストリームは、図3のClip2のIN\_time2に対応するピクチャ以後の符号化ビデオストリームからなる。そして、そのビデオストリームは、正しくデコード開始することができて、これに続くClip2のビデオストリームに接続され、一つの連続でMPEG2規格に従ったエレメンタリーストリームとなるように再エンコードされている。Bridge-Clipを作るためには、一般に、数枚のピクチャは再エンコードしなければならず、それ以外のピクチャはオリジナルのClipからコピーすることができる。

#### 【0071】

図4(a)は、図3に示すClip1をその復号順に示したものであり、プレーヤは、先行するClip1のP5のソースパケット番号 (SPN\_exit\_from\_previous\_Clip) から図4(b)に示すBridge-Clipにジャンプする。ここで、Bridge-Clipの図1に示したD2、即ち、BridgeSequenceにおけるSPN\_arrival\_time\_discontinuityの前までのビデオデータに相当するClip1のOUT\_time1側のストリームデータにおいて、B4までのデータd1は、Clip1をそのままコピーしたデータからなり

、これに続くデータ P 7, B 6 は本来はオリジナルのClip1の B 6 及び B 7 になるが、Clip1を復号し圧縮していない画像データに戻してから再びエンコードして P 7 及び B 6 としたデータ d 2 となっている。また、Bridge-Clipの図 1 に示した D 3、即ち、BridgeSequence におけるSPN\_arrival\_time\_discontinuity以降のビデオデータに相当するClip2のIN\_time2側のストリームデータにおいても、オリジナルのClip2の b 4、p 5、p 8、b 6、b 7 は、Clip2を一旦復号して圧縮していない画像データに戻してから再びエンコードして新たに作成されたデータ (i 0、p 1、p 4、b 2、b 3) d 3 となり、それ以降Clip2のSPN\_enter\_to\_current\_Clipにジャンプするまでの間のデータ d 4 は、Clip2をそのままコピーしたものとなっている。

#### 【0072】

次に、図 3 に示すようなビデオストリーム (Clip1及びClip2) をシームレス接続する場合に、上記第 2 の方法であるBridgeSequenceを使用しないでシームレス接続を実現する例を図 5 に示す。図 5 において、Clip1及びClip 2 は、その復号順序にピクチャを示す。BridgeSequenceを使用しない場合であっても、図 4 に示したBridgeSequenceを使用する場合と同様に、接続点 (コネクションポイント: conection point) 付近のストリームは、一旦復号して圧縮していないデータに戻してから最適なピクチャタイプに再デコードされる。即ち、Clip1 のビデオストリームは、図 3 のOUT\_time1に対応するピクチャまでの符号化ビデオストリームからなり、それは、1つの連続でMPEG2規格に従ったエレメンタリーストリームとなるように、B 6, B 7が再エンコードされたデータ (P 7, B 6) d 5 とされている。同様にして、Clip2 のビデオストリームは、図 3 のClip2のIN\_time2に対応するピクチャ以後の符号化ビデオストリームからなり、それは、1つの連続でMPEG2規格に従ったエレメンタリーストリームとなるように、b 4, p 5, p 8, b 6, b 7が再エンコードされたデータ (i 0, p 1, p 4, b 2, b 3) d 6 とされている。

#### 【0073】

次に、TS1及びTS2のそれぞれの多重化ストリームの符号化制限について説明する。図 6 は、オーディオの表示のオーバーラップを説明する図であって、TS1及

びTS2におけるビデオのプレゼンテーションユニット及びオーディオのプレゼンテーションユニットを示す模式図である。

#### 【0074】

図6に示すように、TS1のオーディオストリームの最後のオーディオフレームA\_endは、TS1の最後の表示ピクチャの表示終了時(OUT\_time1)に等しい表示時刻を持つオーディオサンプルを含んでいる。また、TS2のオーディオストリームの最初のオーディオフレームA\_startは、TS2の最初の表示ピクチャの表示開始時(IN\_time2)に等しい表示時刻を持つオーディオサンプルを含んでいる。これにより、TS1とTS2との接続点において、オーディオプレゼンテーションユニットのシーケンスにギャップは存在せず、2オーディオフレーム区間未満のオーディオプレゼンテーションユニットの長さで定義されるオーディオオーバーラップ(Audio overlap)が生じる。接続点におけるTSは、後述するDVR-STD (Digital Video Recording-System Target Decoder) に従ったDVR MPEG2TSである。

#### 【0075】

DVR-STDは、シームレス接続された2つのPlayItemによって参照されるAVストリームの生成および検証の際におけるデコード処理をモデル化するための概念モデルである。DVR-STDモデル(DVR MPEG2 transport stream player model)を図7に示す。

#### 【0076】

図7に示すように、本実施の形態における情報処理装置(DVR MPEG2 transport stream player model、以下プレーヤという。)1は、シームレスに再生されるよう接続されたTSからトランスポートパケット(Transport packets)を読み出し出力する出力部10と、出力部10からのトランスポートパケットをデコードするデコーダ(DVR-STD)20とから構成される。このデコーダ20は、後述するように、上述した従来のDVR-STDにおけるトランスポートパケットの入力タイミングとオーディオバッファの容量とを変更したものである。出力部10において、読出部(DVRdrive)11から読出しレートRUDで読み出されたTSファイルは、リードバッファ12にバッファリングされ、このリードバッファ12からソースパケット(Source packets)がソースデパケットタイザ部(source depacketi

zer) 13ヘビットレートRMAXで読み出される。RMAXは、ソースパケット  
ストリームのビットレートである。

#### 【0077】

パルス発振器 (27MHz X-tal) 14は、27MHzのパルスを発生する。アラ  
イバルクロックカウンタ (Arrival time clock counter) 15は、この27MHz  
の周波数のパルスをカウントするバイナリカウンタであり、ソースデパケッタイ  
ザ部13に、時刻t(i)におけるArrival time clock counterのカウント値Arriva  
l\_time\_clock(i)を供給する。

#### 【0078】

上述したように、1つのソースパケットは、1つのトランスポートパケットと  
そのarrival\_time\_stampを持つ。現在のソースパケットのarrival\_time\_stamp  
がarrival\_time\_clock(i)のLSB30ビットの値と等しい時、ソースデパケッ  
タイザ部13からそのトランスポートパケットが出力される。TS\_recording\_rat  
eは、TSのビットレートである。

#### 【0079】

また、図7に示すn、TBn、MBn、EBn、TBSys、BSys、Rxn、Rbxn、Rxsys、Dn、D  
sys、On及びPn(k)の表記方法は、ISO/IEC13818-1 (MPEG2 Systems規格)のT-STD  
(ISO/IEC 13818-1で規定されるtransport stream system target decoder)に  
定義されているものと同じである。すなわち、次の通りである。

#### 【0080】

n: エレメンタリーストリームのインデクス番号

TBn: エレメンタリーストリームnのトランスポートバッファ

MBn: エレメンタリーストリームnの多重バッファ (ビデオストリームについての  
み存在)

EBn: エレメンタリーストリームnのエレメンタリーストリームバッファ、ビデオ  
ストリームについてのみ存在

TBSys: 復号中のプログラムのシステム情報のための入力バッファ

BSys: 復号中のプログラムのシステム情報のためのシステムターゲットデコーダ  
内のメインバッファ



Rxn: データがTBnから取り除かれる伝送レート

Rbxn: PESパケットペイロードがMBnから取り除かれる伝送レート (ビデオストリームについてのみ存在)

Rxsys: データがTBsysから取り除かれる伝送レート

Dn: エレメンタリーストリームnのデコーダ

Dsys: 復号中のプログラムのシステム情報に関するデコーダ

On: ビデオストリームnの再配列バッファ (re-ordering buffer)

Pn(k): エレメンタリーストリームnのk番目のプレゼンテーションユニット

#### 【0081】

次に、デコーダ20のデコーディングプロセスについて説明する。まず、1つのDVR MPEG2 TSを再生している時のデコーディングプロセスを説明する。

#### 【0082】

単一のDVR MPEG2TSを再生している間は、トランスポート packets をTB1、TBn又はTBsysのバッファへ入力するタイミングは、ソース packets のarrival\_time\_stampにより決定される。

#### 【0083】

TB1、MB1、EB1、TBn、Bn、TBsys及びTBsysのバッファリング動作の規定は、ISO/IEC 13818-1に規定されているT-STDと同じである。復号動作と表示動作の規定もまた、ISO/IEC 13818-1に規定されているT-STDと同じである。

#### 【0084】

次に、シームレス接続されたPlayItemを再生している間のデコーディングプロセスについて説明する。図8は、あるAVストリーム (TS1) から、これにシームレスに接続された次のAVストリーム (TS2) へと移る時のトランスポート packets の入力、復号、及び表示のタイミングチャートである。

#### 【0085】

ここでは、シームレス接続されたPlayItemによって参照される2つのAVストリームの再生について説明をすることにし、以後の説明では、図1又は図2に示した、シームレス接続されたTS1及びTS2の再生について説明する。従って、TS1は、先行するストリームであり、TS2は、現在のストリームである。

## 【0086】

あるAVストリーム (TS1) からそれにシームレスに接続された次のAVストリーム (TS2) へと移る間には、TS2のアライバルタイムベースの時間軸 (図8においてATC2で示される) は、TS1のアライバルタイムベースの時間軸 (図8におけるATC1で示される) と同じ時間軸でない。また、TS2のシステムタイムベースの時間軸 (図8においてSTC2で示される) は、TS1のシステムタイムベースの時間軸 (図8においてSTC1で示される) と同じ時間軸でない。ビデオの表示は、シームレスに連続していることが要求される。オーディオのプレゼンテーションユニットの表示時間にはオーバーラップがあってもよい。

## 【0087】

ここで、本実施の形態におけるプレーヤ1においては、上記特許文献1乃至3に記載のプレーヤ101に対し、以下の2点を変更することで、オーディオバッファを最適な容量とするものである。まず、1つ目の変更点について説明する。1つ目の変更点は、あるAVストリーム (TS1) からそれにシームレスに接続された次のAVストリーム (TS2) へと移る時、TS1の最後のパケットまでのデコーダ20への入力をそれらソースパケットのarrival\_time\_stampによって決定することとする点である。

## 【0088】

即ち、上述した如く、従来は、TS1の最後のビデオパケットがTB1に入力終了する時刻T1から、TS1の最後のバイトが入力終了するまでの時刻T2までの間は、arrival\_time\_stampを無視して、TSの最大ビットレートで、トランスポートパケットがバッファに入力されていたのに対し、本実施の形態においては、T1からT2までの間のソースパケットの入力を、時刻T1までと同様にして、TS1のソースパケットのarrival\_time\_stampによって決定する。これにより、従来ソースパケットのarrival\_time\_stampを無視して、TSの最大ビットレートRMAXで入力するために必要とされていた1秒分の付加的なバッファは、不要となる。

## 【0089】

この場合のデコーダ20への入力タイミングについて、図8を参照して説明する。

## 【0090】

## (1) 時刻T1までの時間

時刻T1までの時間、即ち、TS1の最後のビデオパッケージがデコーダ20のTB1に入力終了するまでは、デコーダ20のTB1、TBn又はTBsysのバッファへの入力タイミングは、TS1のソースパッケージのarrival\_time\_stampによって決定される。

## 【0091】

## (2) 時刻T1からT2まで

TS1の残りのパッケージがデコーダ20へ入力するタイミングもまた、TS1のソースパッケージのarrival\_time\_stampによって決定される。TS1の最後のバイトがバッファへ入力する時刻は、時刻T2である。

## 【0092】

## (3) 時刻T2以後

T2の時刻において、arrival time clock counterは、TS2の最初のソースパッケージのarrival\_time\_stampの値にリセットされる。デコーダ20のTB1、TBn又はTBsysのバッファへの入力タイミングは、TS2のソースパッケージのarrival\_time\_stampによって決定される。

## 【0093】

即ち、デコーダ20への入力タイミングは、TS1の最後のバイトがバッファへ入力する時刻T2までは、デコーダ20のTB1、TBn又はTBsysのバッファへの入力タイミングをTS1のソースパッケージのarrival\_time\_stampによって決定し、T2以降は、TS2のソースパッケージのarrival\_time\_stampによって決定する。

## 【0094】

次に、ビデオのプレゼンテーションタイミングについて説明する。ビデオプレゼンテーションユニットの表示は、上述の図1又は図2に示すような接続点（コネクションポイント）を通して、ギャップなしに連続しなければならない。即ち、TS1の最後のビデオデータ（第1のピクチャ）に、TS2の最初のビデオデータ（第2のピクチャ）が連続して再生される。ここで、  
STC1: TS1のシステムタイムベースの時間軸（図8ではSTC1と示す。）

STC2: TS2のシステムタイムベースの時間軸(図8ではSTC2と示す。)正確には、STC2は、TS2の最初のPCR((Program Clock Reference))がT-STDに入力した時刻から開始される。)とする。

#### 【0095】

また、STC1とSTC2との間のオフセットは、次のように決定される。即ち、  
 $PTS1_{end}$ : TS1の最後のビデオプレゼンテーションユニットに対応するSTC1上のPTS  
 $PTS2_{start}$ : TS2の最初のビデオプレゼンテーションユニットに対応するSTC2上のPTS  
 $T_{pp}$ : TS1の最後のビデオプレゼンテーションユニットの表示期間  
 とすると、2つのシステムタイムベースの間のオフセット値 $STC\_delta$ は、下記式(6)のように計算される。

#### 【0096】

#### 【数6】

$$STC\_delta = PTS1_{end} + T_{pp} - PTS2_{start} \quad \dots (6)$$

#### 【0097】

次に、オーディオのプレゼンテーションタイミングについて説明する。TS1とTS2との接続点において、オーディオプレゼンテーションユニットの表示タイミングのオーバーラップがあってもよく、それは0以上であって2オーディオフレーム未満である(図8のオーディオオーバーラップを参照)。プレーヤには、どちらのオーディオサンプルを選択するかということ及び、オーディオプレゼンテーションユニットの表示を接続点の後の補正されたタイムベースに再同期する制御が要求される。

#### 【0098】

TS1からそれにシームレスに接続された次のTS2へと移る時、デコーダ20のシステムタイムクロックの制御について、プレーヤが行う処理を説明する。

#### 【0099】

時刻T5において、TS1の最後のオーディオプレゼンテーションユニットが表示される。システムタイムクロックは、時刻T2からT5の間にオーバーラップしていてもよい。この区間では、デコーダ20は、システムタイムクロックを古いタイムベースの値(STC1)を新しいタイムベースの値(STC2)に切り替える。STC2の値は、下記式(7)のように計算できる。

【0100】

【数7】

$$STC2 = STC1 - STC\_delta \quad \dots (7)$$

【0101】

次に、TS1から、このTS1にシームレスに接続される次のTS2へと移る時、TS1及びTS2が満たさなければならない符号化条件を説明する。ここで、  
 $STC1^{1\_end}$ : TS1の最後のパケットの最後のバイトがデコーダ20へ到着する時のシステムタイムベースSTC1上のSTCの値  
 $STC2^{2\_start}$ : TS2の最初のパケットの最初のバイトがデコーダ20へ到着する時のシステムタイムベースSTC2上のSTCの値  
 $STC2^{1\_end}$ :  $STC1^{1\_end}$ の値をシステムタイムベースSTC2上の値に換算した値

とすると、 $STC2^{1\_end}$ は、下記式(8)のようにして計算される。

【0102】

【数8】

$$STC2^{1\_end} = STC1^{1\_end} - STC\_delta \quad \dots (8)$$

【0103】

ここで、デコーダ20がDVR-STDに従うために、次の2つの条件を満たすことが要求される。

【0104】

(条件1)

TS2の最初のパケットのデコーダ20への到着タイミングは、次に示す不等式(9)を満たさなければならない。

【0105】

【数9】

$$STC2^2_{start} > STC2^1_{end} \quad \dots (9)$$

【0106】

この上記不等式(9)を満たすように、Clip1及び／又はClip2の部分的なストリームを再エンコード及び／又は再多重化することが必要になる。

【0107】

(条件2)

STC1とSTC2とを同じ時間軸上に換算したシステムタイムベースの時間軸上において、TS1からのビデオパケットの入力とそれに続くTS2からのビデオパケットの入力は、ビデオバッファをオーバーフロー及びアンダーフローさせてはならない。また、STC1とSTC2とを同じ時間軸上に換算したシステムタイムベースの時間軸上において、TS1からのパケットの入力とそれに続くTS2からのパケットの入力は、デコーダ20のすべてのバッファをオーバーフローおよびアンダーフローさせてはならない。

【0108】

図9は、あるAVストリーム(TS1)からそれにシームレスに接続された次のAVストリーム(TS2)へと移る時のトランスポートパケットの入力、復号、表示のタイミングチャートの他の例である。この場合もまた、TS1の最後のパケットまでのデコーダ20への入力をそれらソースパケットのarrival\_time\_stampによって決定することは同様であるが、図8に示すタイミングチャートとの一点の違いとして、図9に示すように、TS1の最後のパケットの直後にTS2の最初のパケット入力する必要がないように、所定の時間間隔(delta1:時刻T2~T2'の間)を設けている。これにより、TS2の最初のパケットの入力タイミングの決定が図8の場合よりも柔軟であるので、TS2の符号化を容易にする効果がある。

## 【0109】

この場合のデコーダ20への入力タイミングについて、図9を参照して説明する。

## 【0110】

## (1) 時刻T2までの時間

時刻T2までの時間、すなわち、TS1の最後のパケットの最後のバイトがデコーダ20へ入力終了するまでは、デコーダ20のTB1、TBnまたはTBsysのバッファへの入力タイミングは、TS1のソースパケットのarrival\_time\_stampによって決定される。

## 【0111】

## (2) 時刻T2' 以後

時刻T2からdelta1の時間の後、時刻T2'の時刻において、arrival time clock counterは、TS2の最初のソースパケットのarrival\_time\_stampの値にリセットされる。デコーダ20のTB1、TBn又はTBsysのバッファへの入力タイミングは、TS2のソースパケットのarrival\_time\_stampによって決定される。

## 【0112】

ここで、図9に示すように、delta1を設ける場合、上述の $STC2^2_{start}$ と $TC2^1_{end}$ は下記の関係式(10)を満たさなければならない。

## 【0113】

## 【数10】

$$STC2^2_{start} > TC2^1_{end} + \delta 1 \quad \dots (10)$$

## 【0114】

以上説明したように、デコーダ20への入力タイミングを、先行するTS1の最後のビデオパケットを入力した後もソースパケットのアライバルタイムスタンプに従って行う点が1つ目の変更点である。次に2つ目の変更点について説明する。

## 【0115】

2つ目の変更点としては、デコーダ20のオーディオバッファのサイズを、次

の条件が満たされるような十分な大きさに変更する。この条件とは、TS1からこれにシームレスに接続される次のTS2へと移る時、TS1の最後のトランスポートパッケージを入力後に、TS2の最初にデコードされるピクチャ（Iピクチャ）をそのデコードタイミングまでにビデオバッファへ入力できることである。

#### 【0116】

この条件を満たすために、オーディオバッファの必要な容量の最大値は次のような値である。即ち、「Iピクチャの最大のビット量をそのデコードタイミングまでにビデオバッファへ入力できる時間」に相当する長さのオーディオのデータ量を蓄えられる大きさである。オーディオバッファの必要量の最大値 $EBn\_max$ は、次の式（11）で計算できる。

#### 【0117】

#### 【数11】

$$EBn\_max = (I\_max / Rv) * Ra \text{ [bits]} \quad \dots (11)$$

#### 【0118】

ここで、 $I\_max$ は、Iピクチャの最大のビット量であり、これは、図7に示すビデオコードバッファEB1の大きさである。 $Rv$ は、ビデオコードバッファEB1への入力ビットレートである。上記式（11）に示すように、求めるべきオーディオバッファの大きさ $EBn\_max$ は、ビデオエレメンタリーストリームバッファ（EB1）への入力ビットレートでビデオコードバッファEB1のバッファ占有量をゼロから $I\_max$ にするまでにかかる時間（ $I\_max / Rv$ ）に $Ra$ を掛けた値である。

#### 【0119】

また、具体的な値として、少なくとも100ミリ秒分のオーディオデータを蓄えられるバッファサイズを推奨する。これは以下の理由による。即ち、Iピクチャを0.5秒ごとに符号化する場合、Iピクチャのビットサイズは、一般的に、符号化ビットレートの10%以下である。例えば、符号化ビットレートが10Mbpsの場合、Iピクチャのサイズは、通常1Mbits以下である。

#### 【0120】



従って、第1に、少なくとも、100ミリ秒分の時間があれば、Iピクチャをそのデコードタイミングまでにデコーダ20のビデオバッファへ入力できる。また、第2にデコーダ20のオーディオバッファが、100ミリ秒分のオーディオデータを蓄えられるのであれば、オーディオデータをそれが再生されるタイミングよりも100ミリ秒だけ早く、オーディオバッファに入力終了するようにTS1を多重化できる。従って、オーディオバッファを、少なくとも100ミリ秒分のオーディオデータを蓄えられるバッファサイズとすれば、上記第1及び第2の理由により、TS1からそれにシームレスに接続される次のTS2へと移る時、TS1の最後のトランスポート packets を入力後に、TS2の最初にデコードされるピクチャ（Iピクチャ）をビデオバッファへ入力終了する（Iピクチャのコードタイミング）までの時間として少なくとも100ミリ秒分を確保することができる。

#### 【0121】

100ミリ秒分のオーディオデータを蓄えられるオーディオバッファの容量を具体的に計算すると次のようになる。

#### 【0122】

640 kbpsのドルビーAC3オーディオストリームの場合： $640 \text{ kbps} \times 0.1 \text{ s}$   
 $\text{ec} = 64 \text{ kbitS} = 8 \text{ kBytes}$

Linear PCMオーディオストリーム、24bitSample、96 kHz sampling frequency、8 channelsの場合： $(24 \text{ bitSample} \times 96000 \text{ samples/sec} \times 8 \text{ ch}) \times 0.1 \text{ sec} =$   
230400 Bytes

#### 【0123】

次に、以上説明した本実施の形態のデコーダ20のように、DVR-STDのオーディオバッファのサイズを、「TS2の最初にデコードされるピクチャ（Iピクチャ）をそのデコードタイミングまでにビデオバッファへ入力できる時間に相当するオーディオのデータ量を蓄えられる大きさ」に変更することによる効果を図10及び図11を用いて更に詳細に説明する。

#### 【0124】

例として、オーディオストリームが、ビットレート640 kbps、サンプリング周波数48 kHzのAC3オーディオストリームの場合を説明する。AC3オーディオス

トリームの1オーディオフレームのサンプル数は1536 samplesであるので、その時間長は、32ミリ秒である。また、1オーディオフレームのバイトサイズは、2560 Bytesである。

#### 【0125】

図10(a)及び(b)は、従来のDVR-STDの場合であり、オーディオのバッファサイズが4 kBytesである場合に、TS1からそれにシームレスに接続される次のTS2へと移る時のDVR-STDの夫々ビデオバッファ及びオーディオバッファのビット占有量の変化の例を示すグラフ図である。この図10において、TS1のビデオ／オーディオデータのバッファ遷移を破線で示し、また、TS2のビデオ／オーディオデータのバッファ遷移を実線で示す。

#### 【0126】

4 kBytesのオーディオバッファは、50ミリ秒分のオーディオデータを蓄えられる。従って、TS1の最後のオーディオパケットの最後のバイトがDVR-STDに到着する時刻であるSTC1<sup>1</sup> audio\_endにおいて、オーディオデータをそれが再生されるタイミングよりも50ミリ秒だけ早く入力終了するようにTS1を多重化できる。しかしながら、50ミリ秒は、TS2の最初にデコードされるピクチャ(Iピクチャ)をそのデコードタイミングまでにビデオバッファへ入力するためには、不十分である。そのため、TS2の最初にデコードされるピクチャ(Iピクチャ)のサイズを小さくするようにエンコードを制限することになり、画質が悪くなる問題がある。

#### 【0127】

即ち、4 kBytesのオーディオバッファでは、先送りすることができるオーディオデータは50ミリ秒であるので、TS2の最初のIピクチャをビデオバッファへ入力するための時間である図10に示すスタートアップディレイ(start-up delay) t<sub>1</sub>は、最大50ミリ秒と小さくなる。従って、TS2の最初のIピクチャを入力する時間を十分とれず、IピクチャサイズS<sub>1</sub>が小さくなり、符号化が制限されIピクチャの画質が劣化する。なお、このstart-up delayを大きくするため、上述した如く、従来は、4 kBytesに加えて1秒分程度の付加的なバッファを設け、且つT<sub>1</sub>～T<sub>2</sub>の間はTSの最大レートR<sub>MAX</sub>で入力する必要があった。

ここでは、ビットレート640 kbpsのAC3オーディオストリームについて説明しているが、上述した如く、マルチチャネルのLPCMオーディオに対応可能なように設計すると、1秒分程度の付加的なバッファは極めて大きくなってしまう。

#### 【0128】

この問題を解決するために、本実施の形態におけるデコーダ20のように、DVR-STDのオーディオバッファサイズを例えば8 kBytesに変更する。図11 (a) 及び (b) は、オーディオバッファの容量を最適化した例を示すものであって、オーディオのバッファサイズが8 kBytesである場合に、TS1からこれにシームレスに接続される次のTS2へと移る時の本実施の形態におけるDVR-STDの夫々ビデオバッファ及びオーディオバッファのビット占有量の変化の例を示すグラフ図である。この図11において、TS1のビデオ／オーディオデータのバッファ遷移を破線で示し、また、TS2のビデオ／オーディオデータのバッファ遷移を実線で示す。

。

#### 【0129】

8 kBytesのオーディオバッファは、100ミリ秒分のオーディオデータを蓄えられる。したがって、TS1の最後のオーディオパケットの最後のバイトがDVR-STDに到着する時刻であるSTC1<sup>1</sup> audio\_endにおいて、オーディオデータをそれが再生されるタイミングよりも100ミリ秒だけ早く入力終了するようにTS1を多重化できる。少なくとも、100ミリ秒あれば、TS2の最初にデコードされるピクチャ (Iピクチャ) をそのデコードタイミングまでにビデオバッファへ入力する余裕ができる。即ち、TS2の最初のIピクチャを入力する時間 (スタートアップディレイ) t<sub>2</sub>を十分とることができ、そのため、TS2の最初にデコードされるピクチャ (Iピクチャ) サイズS<sub>2</sub>を大きくすることができ、従ってIピクチャのエンコード制限が小さく、高画質にすることができる。

#### 【0130】

また、図7に示すようなプレーヤモデル1において、トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、そのアライバルタイムスタンプに基づきデコーダにより読み出されてデコードされるTSは、多重化装置 (情報処理装置) において生成され記録され

たものとしてすることができる。

#### 【0131】

多重化装置は、例えば上述の図3乃至5を参照して説明したように、所定のピクチャで表示終了するよう、再符号化したClip1（第1のビデオ符号化ストリーム）を生成し、このピクチャに続けて表示され、且つ表示開始できるよう再符号化したClip2（第2のビデオ符号化ストリーム）を生成するビデオ符号化部と、Clip1とこのClip1に同期したオーディオ符号化ストリームとを多重化してTS1を生成し、Clip2とこのClip2に同期したオーディオ符号化ストリームとを多重化してTS2を生成する多重化部と、TS1及びTS2からなる多重化ストリームを記録する記録部とを備える。ここで、多重化部においては、上記第2のピクチャであるIピクチャをデコーダ20のビデオバッファに入力するために要する時間分のオーディオデータを当該Iピクチャがデコーダ20に入力開始される前までに入力終了可能なように多重化する。なお、図4に示すように、符号化部においてBridge-Clipを生成し、多重化部においてBridge-Clipも合わせて多重化するようにしてもよいことはもちろんである。

#### 【0132】

このような多重化装置により生成された多重化ストリームが記録された記録媒体には、第1のピクチャで表示終了するTS1と、この第1のピクチャに続けて再生する第2のピクチャから表示開始するTS2とから生成され、TS1及びTS2が夫々のアライバルタイムスタンプに基づいてデコーダ20に入力可能であって、且つ、第2のピクチャであるTS2の最初のピクチャをデコーダに入力するために要する時間分のオーディオデータを第2のピクチャがデコーダ20に入力開始される前までに入力終了可能なように多重化された多重化ストリームが記録されたものとなっている。

#### 【0133】

このように構成された本実施の形態においては、シームレスに接続されたTS1及びTS2を再生する際、TS1の最後のビデオパッケージがデコーダ20のTB1に入力終了した後からTS1の残りのパッケージがデコーダ20へ入力するまでにおいても、トランスポートパッケージをアライバルタイムスタンプに従って入力するように

し、オーディオバッファのサイズを、従来のDVR-STDにおける4 kBytesから、Iピクチャの最大のビット量をそのデコードタイミングまでにビデオバッファへ入力できる時間に相当する長さのオーディオのデータ量を蓄えられる大きさに変更することにより、TS1の最後の packets が入力終了してからTS2の最初のピクチャであるIピクチャをそのデコードタイミングまでに入力する時間（スタートアップディレイ）を十分確保することができるので、Iピクチャの符号化制限を小さくし、高画質とすることができる。

#### 【0134】

また、従来のように付加的なバッファを設ける方法であると、例えばTSにおけるオーディオデータをマルチチャンネルのLPCMオーディオデータとした場合に極めて大きい容量の付加的なバッファが必要になるのに対し、本実施の形態においては、オーディオバッファの容量を上記のように変更し、アライバルタイムスタンプに従ってトランスポート packets を入力するようにすることで、従来必要となっていた付加的なバッファを不要とすることができる。

#### 【0135】

##### 【発明の効果】

以上詳述した本発明においては、トランスポート packets とそのアライバルタイムスタンプとを有するソース packets を単位とするデータ列からなり、第1の多重化ストリームの最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続された多重化ストリームを復号する情報処理装置において、上記多重化ストリームの上記アライバルタイムスタンプに従って上記ソース packets を出力し、オーディオバッファの容量を、上記第2のピクチャをビデオバッファに入力するために要する時間分のオーディオデータをバッファリング可能なものとしたので、ビデオストリームとオーディオストリームの多重化ストリームをビデオフレーム精度で編集して、編集点をシームレスに再生することが可能であり、従来ソース packets のarrival\_time\_stampを無視して、TSの最大ビットレートで入力するために必要とされていた1秒分の付加的なバッファを不要とし、従来よりもデコーダに必要なバッファ量を小さくできると共に、オーディオバッファのサイズを、上記第

2のピクチャを上記ビデオバッファに入力するために要する時間分のオーディオデータをバッファリング可能な容量とすることにより第2のピクチャのエンコード制限が小さく、高画質とすることができる。

#### 【図面の簡単な説明】

##### 【図1】

Bridge-Clipを使用する場合の先行するPlayItemと現在のPlayItemとの関係を示す模式図である。

##### 【図2】

Bridge-Clipを使用しない場合の先行するPlayItemと現在のPlayItemの関係を示す模式図である。

##### 【図3】

ビデオストリームとしてのClip1とClip2とをシームレス接続する例をピクチャの表示順序 (Presentation order) で示す模式図である。

##### 【図4】

図3に示ビデオストリーム (Clip1及びClip2) をシームレス接続する場合に、第1の方法であるBridgeSequenceを使用してシームレス接続を実現する各AVストリームにおけるデータ列を示す模式図である。

##### 【図5】

図3示すビデオストリーム (Clip1及びClip2) をシームレス接続する場合に、第2の方法であるBridgeSequenceを使用しないでシームレス接続を実現する各AVストリームにおけるデータ列を示す模式図である。

##### 【図6】

オーディオの表示のオーバーラップを説明する図であって、TS1及びTS2におけるビデオのプレゼンテーションユニット及びオーディオのプレゼンテーションユニットを示す模式図である。

##### 【図7】

本発明の実施の形態における情報処理装置を示すブロック図である。

##### 【図8】

あるAVストリーム (TS1) から、これにシームレスに接続された次のAVストリ

ーム (TS2) へと移る時のトランスポートパケットの入力、復号、及び表示のタイミングチャートである。

#### 【図 9】

あるAVストリーム (TS1) からそれにシームレスに接続された次のAVストリーム (TS2) へと移る時のトランスポートパケットの入力、復号、表示の他の例を示すタイミングチャートである。

#### 【図 10】

(a) 及び (b) は、従来のDVR-STDの場合であり、オーディオのバッファサイズが 4 kBytesである場合に、TS1からそれにシームレスに接続される次のTS2へと移る時のDVR-STDの夫々ビデオバッファ及びオーディオバッファのビット占有量の変化の例を示すグラフ図である。

#### 【図 11】

(a) 及び (b) は、本発明の実施の形態における効果を説明する図であって、オーディオのバッファサイズが 8 kBytesである場合に、TS1からそれにシームレスに接続される次のTS2へと移る時のDVR-STDの夫々ビデオバッファ及びオーディオバッファのビット占有量の変化の例を示すグラフ図である。

#### 【図 12】

従来の情報処理装置を示すブロック図である。

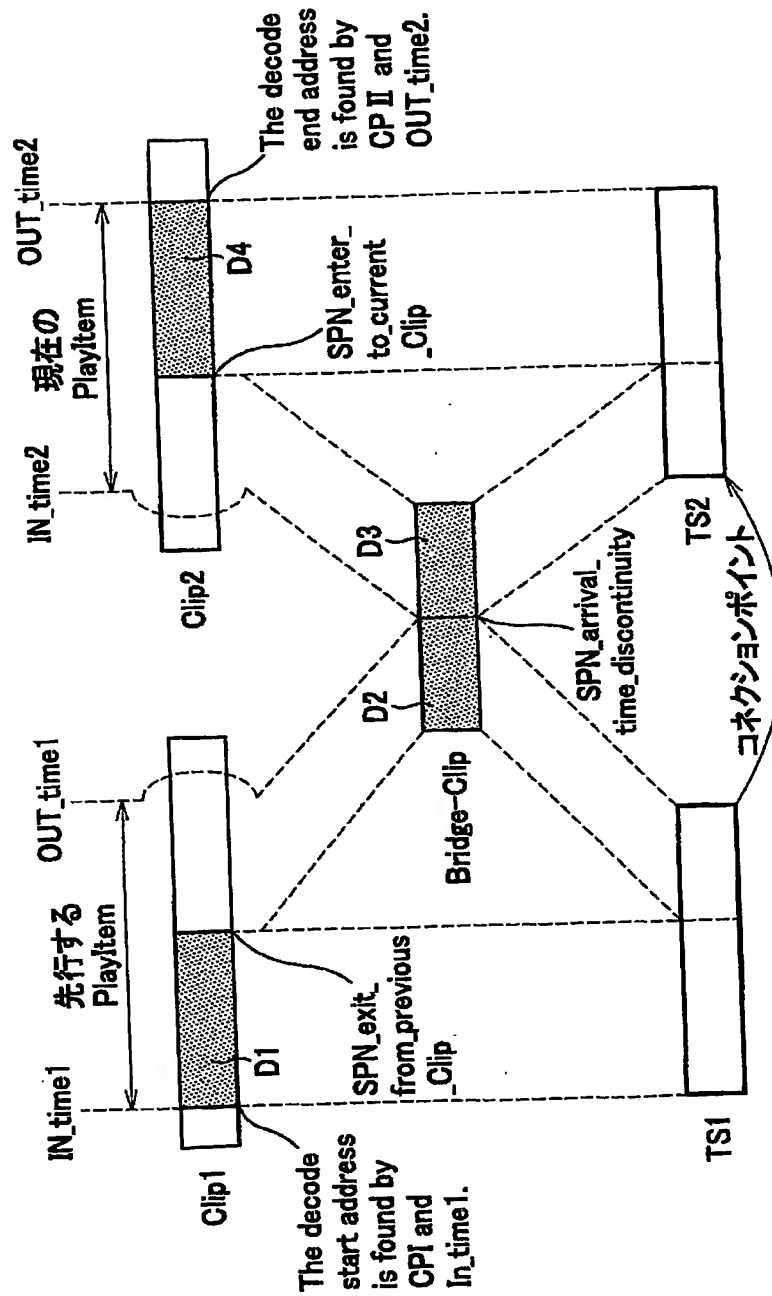
#### 【符号の説明】

11 読出部、12 リードバッファ、13 ソースデパケットタイザ部、14 パルス発振器、15 アライバルタイムクロックカウンタ

【書類名】

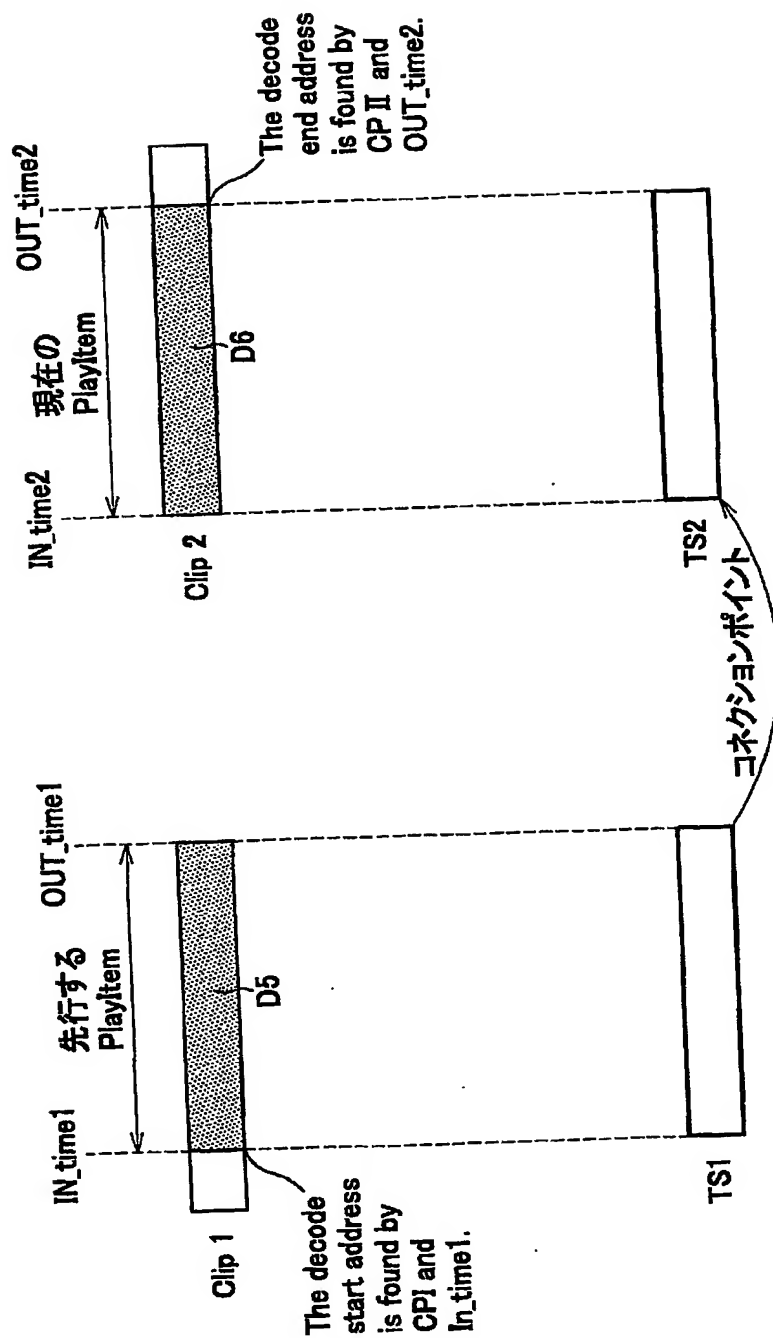
図面

【図1】

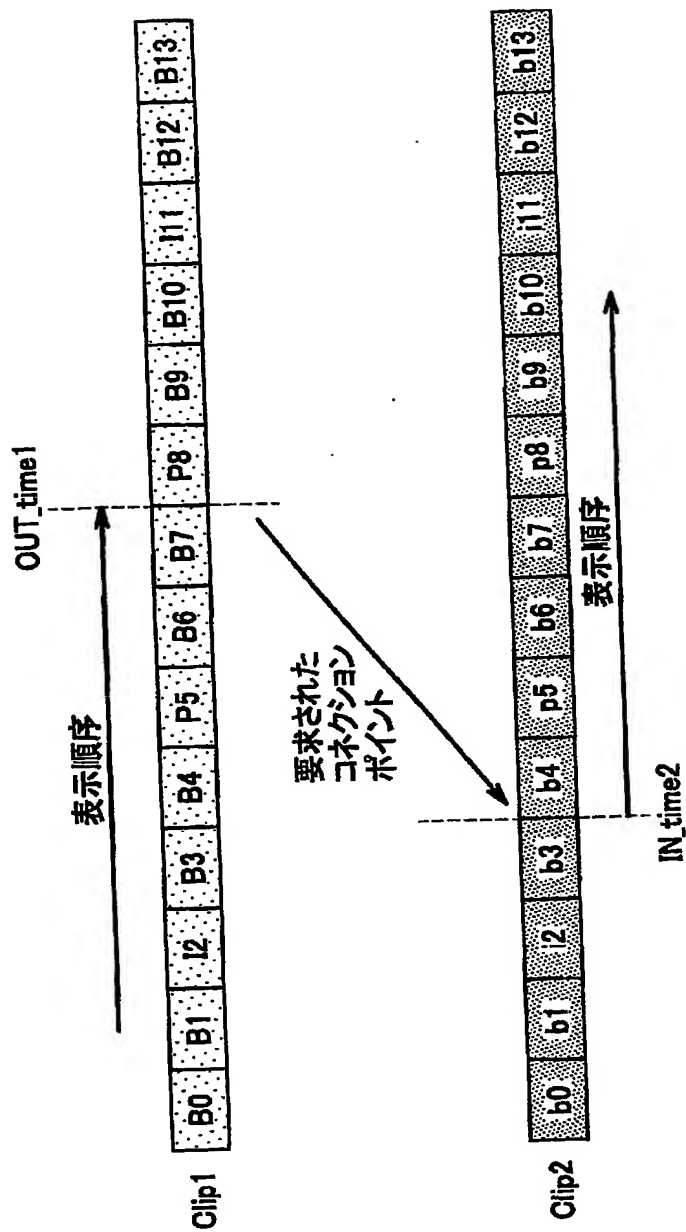




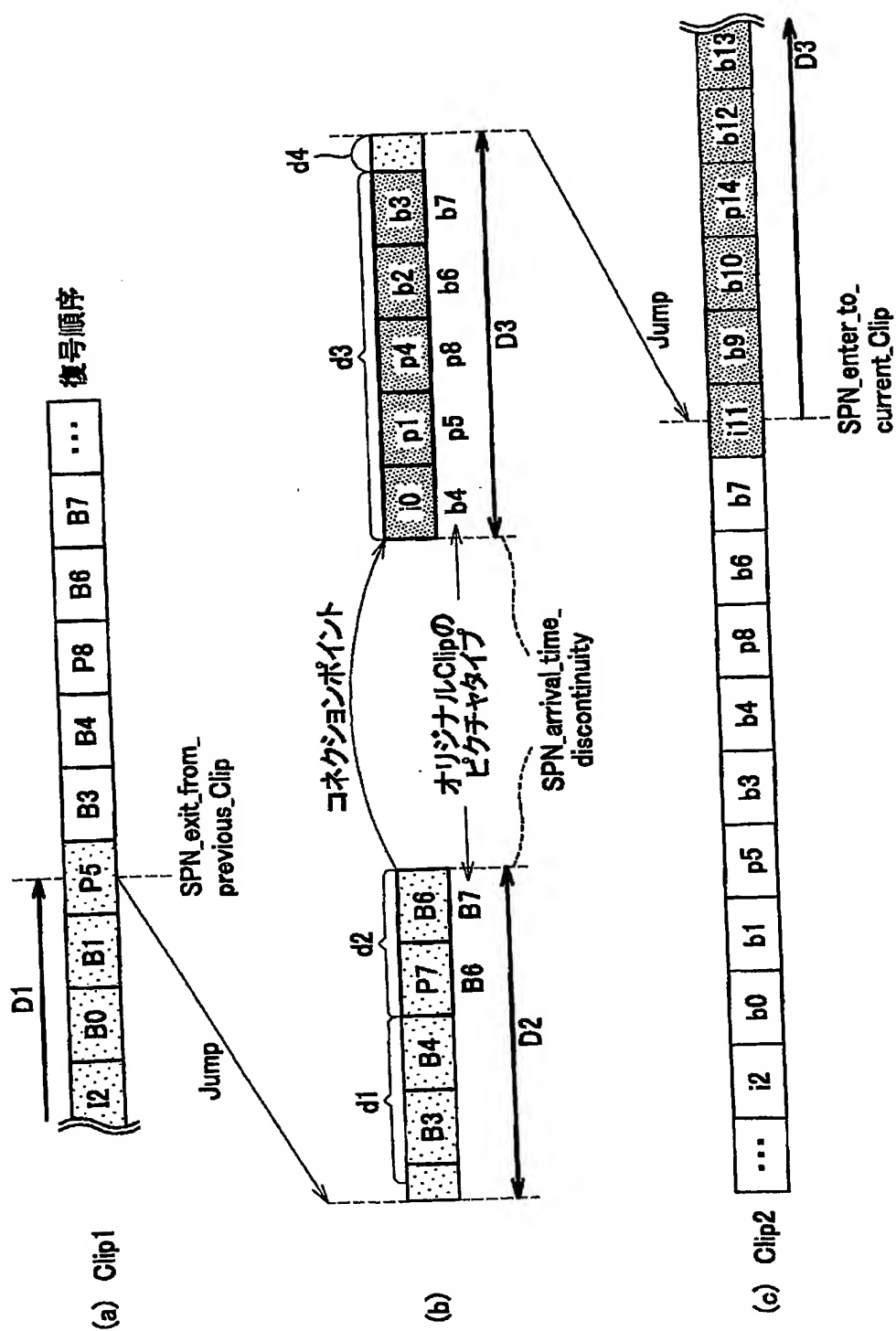
【図2】



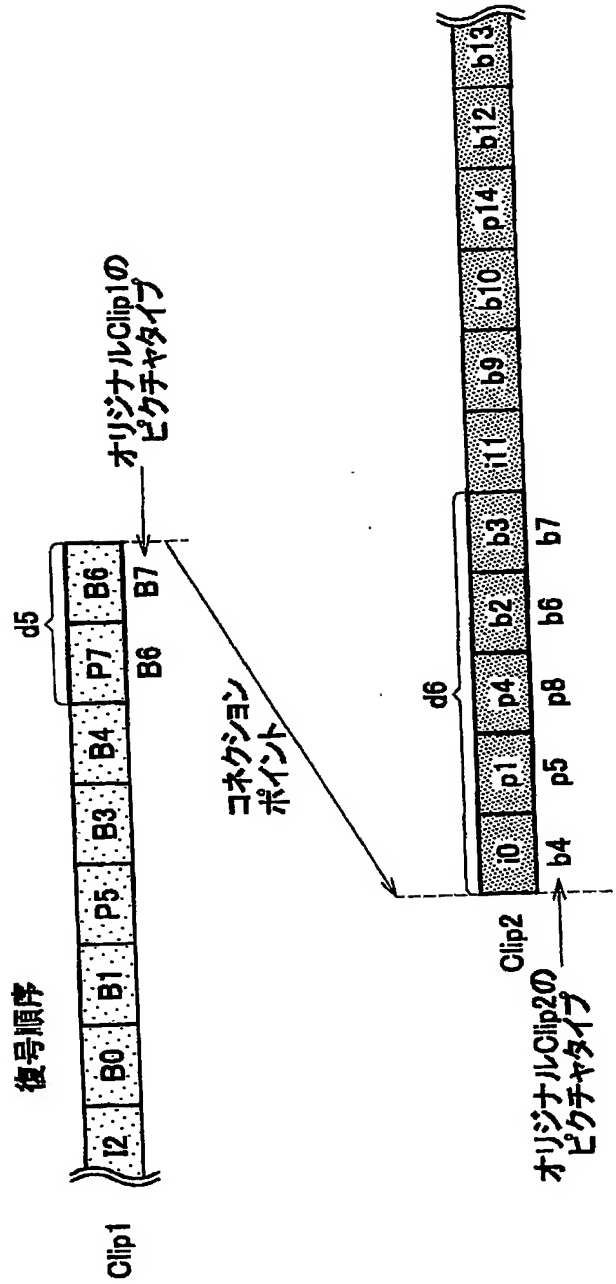
【図3】



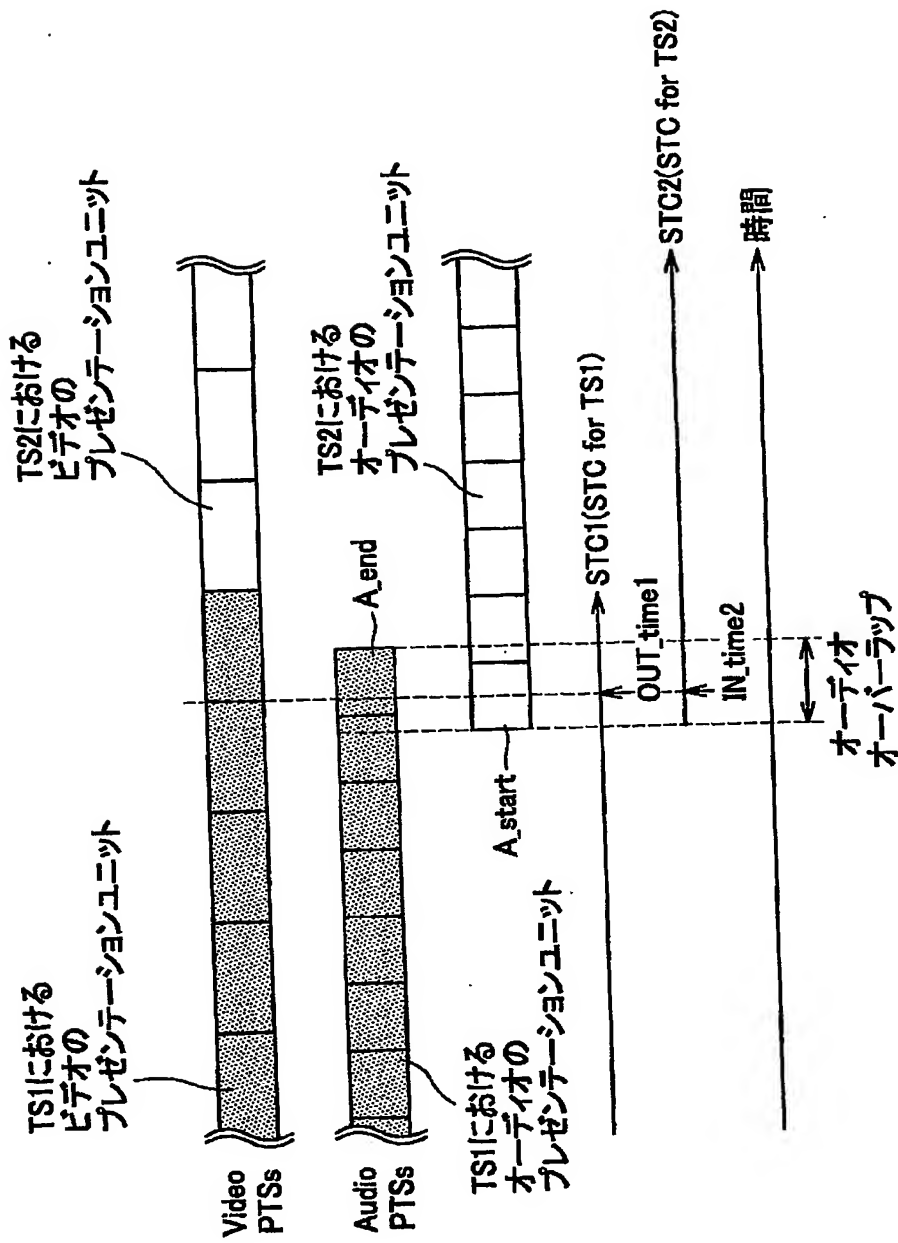
【図4】



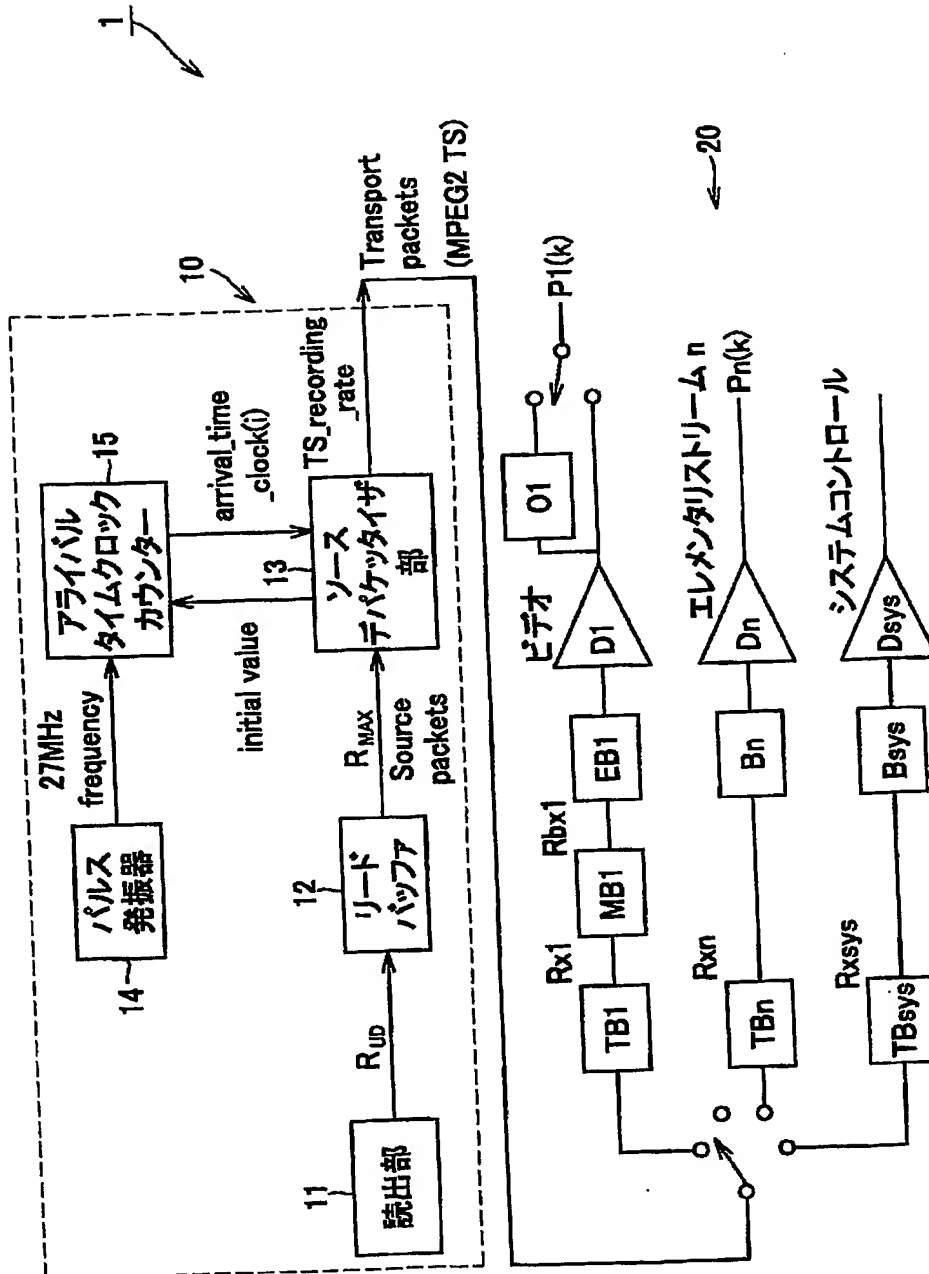
【図5】



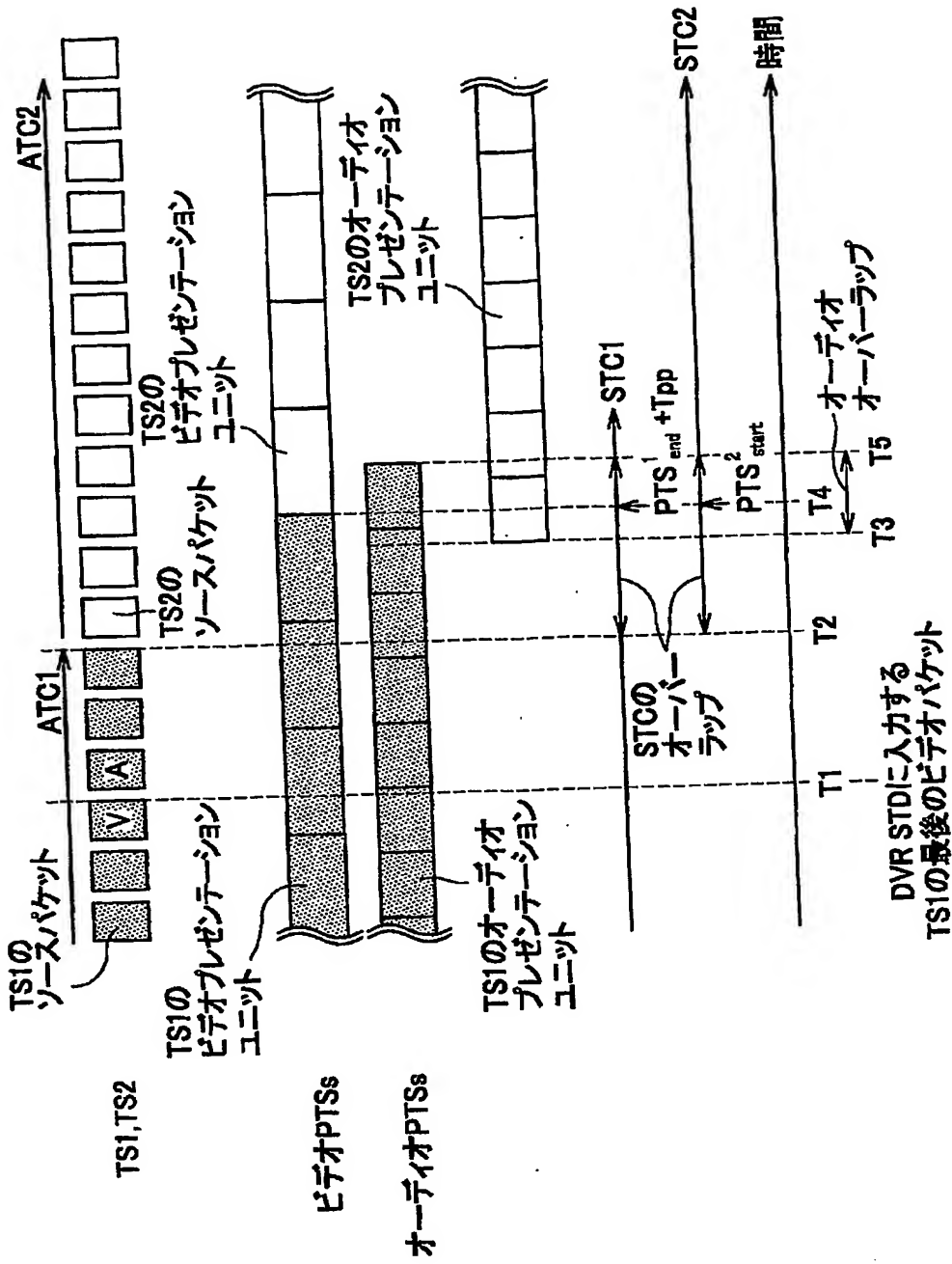
【図 6】



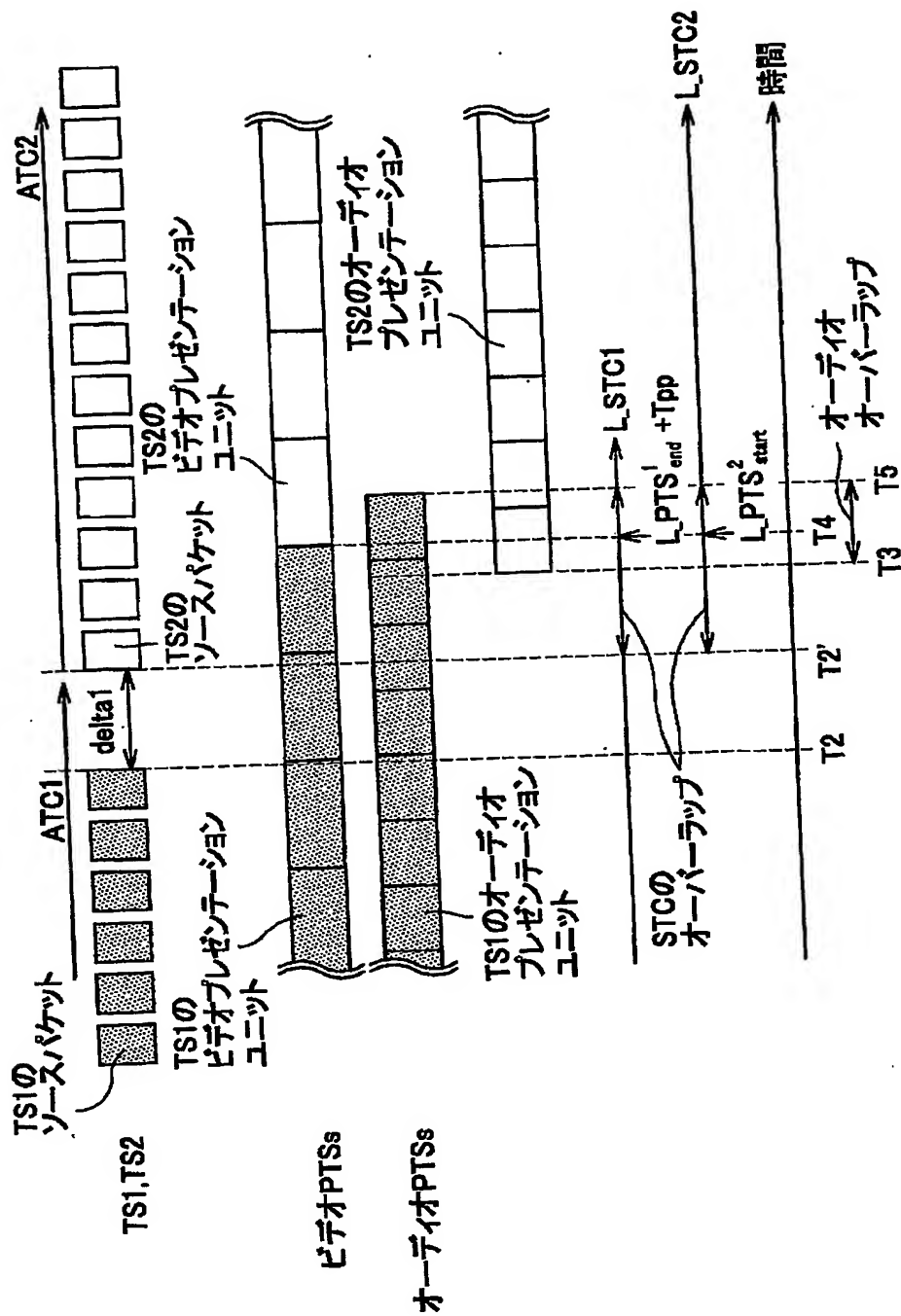
【図7】



【図8】



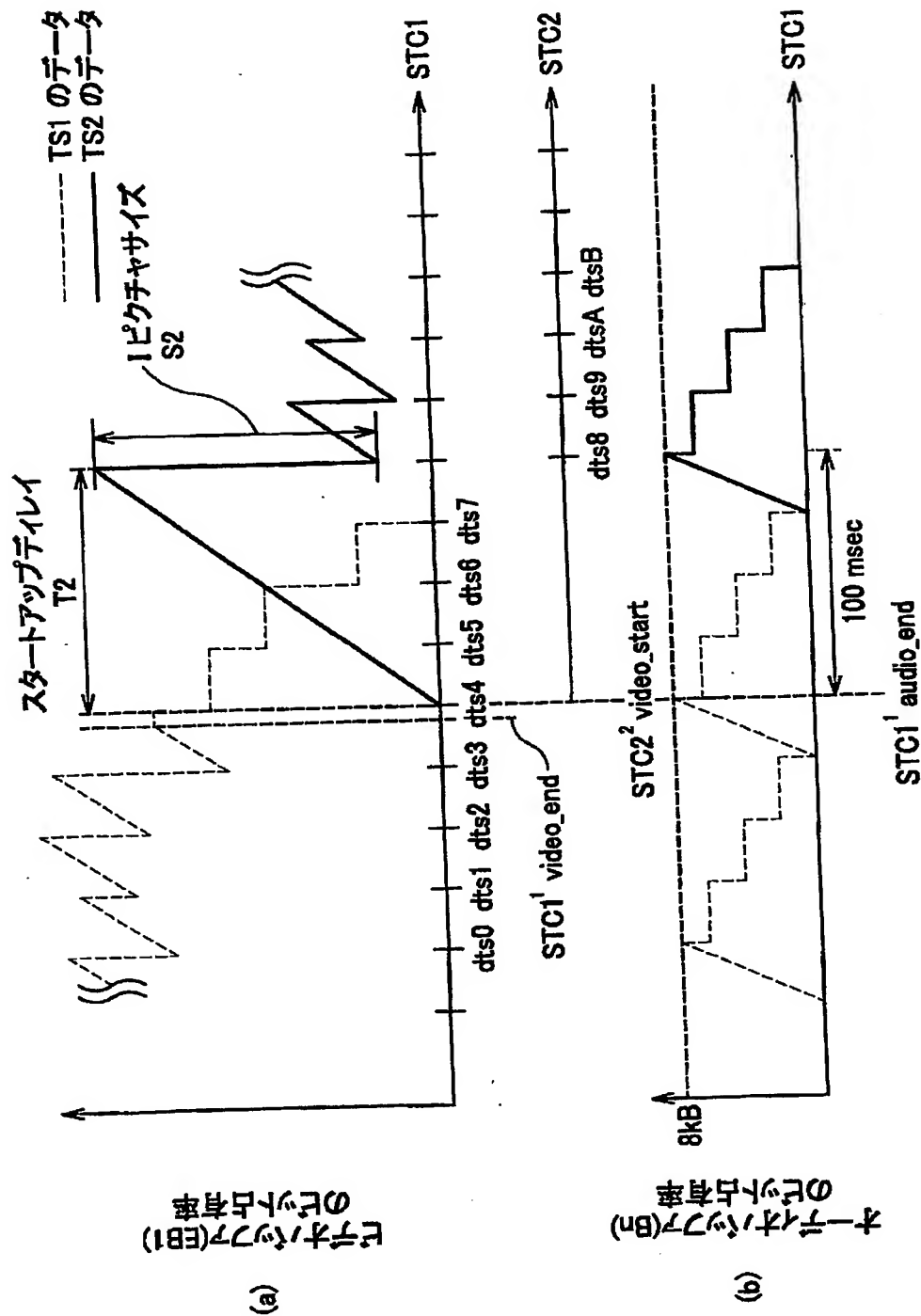
【図9】



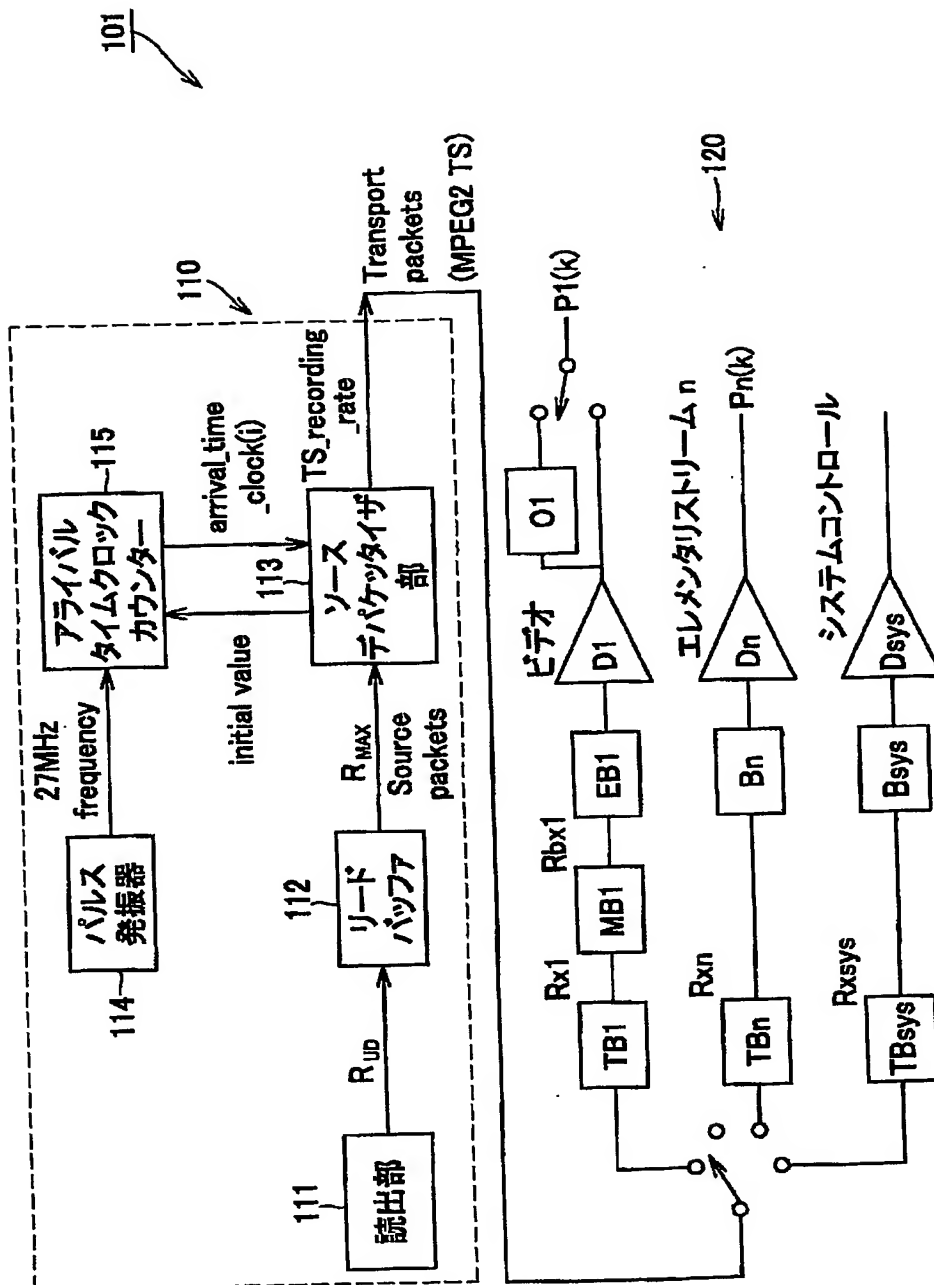




【図11】



【図12】



【書類名】 要約書

【要約】

【課題】 オーディオストリームとビデオストリームとが多重化された2つの多重化ストリームをシームレスに連続して復号する際に、最適な容量のオーディオバッファとする。

【解決手段】 プレーヤモデル1は、シームレスに接続されたTS1及びTS2から、トランスポートパケット及びアライバルタイムスタンプを有するソースパケットを読み出し、そのアライバルタイムスタンプに従ってトランスポートパケットを出力する出力部10と、トランスポートパケットをデコードするデコーダ20とから構成される。出力部10は、ソースパケットのアライバルタイムスタンプに従ってトランスポートパケットをデコーダ20に入力するものとし、デコーダ20は、そのオーディオバッファTBnを、TS2の最初のピクチャであるIピクチャをビデオバッファTB1に入力するために要する時間分のオーディオデータをバッファリング可能な容量とする。

【選択図】 図7

特願 2003-130661

出願人履歴情報

識別番号

[000002185]

1. 変更年月日

[変更理由]

住所  
氏名

1990年 8月30日

新規登録

東京都品川区北品川6丁目7番35号

ソニー株式会社

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**